

FIG.1

2/62

## hCEA-LTA Nucleotide Sequence

```

1   ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCAAGCTCAC TATTGAATCC ACGCCGTTCA ATGTCGCAGA GGGGAAGGAG
151 GTGCTTCTAC TTGTCCACAA TCTGCCCCAG CATCTTTTGT GCTACAGCTG
201 GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA GGATATGTAA
251 TAGGAACTCA ACAAGCTACC CCAGGGCCCCG CATACAGTGG TCGAGAGATA
301 ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC
351 AGGATTCTAC ACCCTACACG TCATAAAGTC AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTCCATCTCC
451 AGCAACAACCT CCAAACCCGT GGAGGACAAG GATGCTGTGG CTTTCACCTG
501 TGAACCTGAG ACTCAGGACG CAACCTACCT GTGGTGGGTA AACAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CAGGACCCTC
601 ACTCTATTCA ATGTCACAAG AAATGACACA GCAAGCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGCCA GGCGCAGTGA TTCAGTCATC CTGAATGTCC
701 TCTATGGCCC GGATGCCCCC ACCATTTCCC CTCTAAACAC ATCTTACAGA
751 TCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA ACCCACCTGC
801 ACAGTACTCT TGGTTTGTCA ATGGGACTTT CCAGCAATCC ACCCAAGAGC
851 TCTTTATCCC CAACATCACT GTGAATAATA GTGGATCCTA TACGTGCCAA
901 GCCCATAACT CAGACACTGG CCTCAATAGG ACCACAGTCA CGACGATCAC
951 AGTCTATGCA GAGCCACCCA AACCCTTCAT CACCAGCAAC AACTCCAACC
1001 CCGTGGAGGA TGAGGATGCT GTAGCCTTAA CCTGTGAACC TGAGATTGAG
1051 AACACAACCT ACCTGTGGTG GGTAATAAAT CAGAGCCTCC CGGTGAGTCC
1101 CAGGCTGCAG CTGTCCAATG ACAAACAGGAC CCTCACTCTA CTCAGTGTCA
1151 CAAGGAATGA TGTAGGACCC TATGAGTGTG GAATCCAGAA CGAATTAAGT
1201 GTTGACCACA GCGACCCAGT CATCCTGAAT GTCCTCTATG GCCCAGACGA
1251 CCCCACCATT TCCCCCTCAT ACACCTATTA CCGTCCAGGG GTGAACCTCA
1301 GCCTCTCCTG CCATGCAGCC TCTAACCCAC CTGCACAGTA TTCTTGGCTG
1351 ATTGATGGGA ACATCCAGCA ACACACACAA GAGCTCTTTA TCTCCAACAT
1401 CACTGAGAAG AACAGCGGAC TCTATACCTG CCAGGCCAAT AACTCAGCCA
1451 GTGGCCACAG CAGGACTACA GTCAAGACAA TCACAGTCTC TGCGGAGCTG
1501 CCCAAGCCCT CCATCTCCAG CAACAACCTC AAACCCGTGG AGGACAAGGA
1515 TGCTGTGGCC TTCACCTGTG AACCTGAGGC TCAGAACACA ACTTACCTGT
1601 GGTGGGTAAA TGGTCAGAGC CTCCCAGTCA GTCCCAGGCT GCAGCTGTCC
1651 AATGGCAACA GGACCCTCAC TCTATTCAAT GTCACAAGAA ATGACGCAAG
1701 AGCCTATGTA TGTGGAATCC AGAACTCAGT GAGTGCAAAC CGCAGTGACC
1751 CAGTACCCTT GGATGTCCTC TATGGGCCGG ACACCCCCAT CATTTCCCCC
1801 CCAGACTCGT CTTACCTTTC GGGAGCGAAC CTCAACCTCT CCTGCCACTC
1851 GGCCTCTAAC CCATCCCCGC AGTATTCTTG GCGTATCAAT GGGATACCGC
1901 AGCAACACAC ACAAGTTCTC TTTATCGCCA AAATCACGCC AAATAATAAC
1951 GGGACCTATG CCTGTTTTGT CTCTAACTTG GCTACTGGCC GCAATAATTC

```

FIG.2A-1

3/62

```
2001 CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACTCTA GTTAATGGCG
2051 ACAAATTATA CCGTGCTGAC TCTAGACCCC CAGATGAAAT AAAACGTTCC
2101 GGAGGTCTTA TGCCCAGAGG GCATAATGAG TACTTCGATA GAGGAACTCA
2151 AATGAATATT AATCTTTATG ATCACGCGAG AGGAACACAA ACCGGCTTTG
2201 TCAGATATGA TGACGGATAT GTTTCCACTT CTCTTAGTTT GAGAAGTGCT
2251 CACTTAGCAG GACAGTCTAT ATTATCAGGA TATTCCACTT ACTATATATA
2301 TGTTATAGCG ACAGCACCAA ATATGTTTAA TGTTAATGAT GTATTAGGCG
2351 TATACAGCCC TCACCCATAT GAACAGGAGG TTTCTGCGTT AGGTGGAATA
2401 CCATATTCTC AGATATATGG ATGGTATCGT GTTAATTTTG GTGTAATTGA
2451 TGAACGATTA CATCGTAACA GGAATATAG AGACCGGTAT TACAGAAATC
2501 TGAATATAGC TCCGGCAGAG GATGGTTACA GATTAGCAGG TTTCCCACCG
2551 GATCACCAAG CTTGGAGAGA AGAACCCTGG ATTCATCATG CACCACAAGG
2601 TTGTGGAAAT TCATCAAGAA CAATTACAGA TGATACTTGT AATGAGGAGA
2651 CCCAGAATCT GAGCACAATA TATCTCAGGA AATATCAATC AAAAGTTAAG
2701 AGGCAGATAT TTTCAGACTA TCAGTCAGAG GTTGACATAT ATAACAGAAT
TCGGGATGAA TTATGA (SEQ ID NO:7)
```

FIG.2A-2

4/62

## CEA-LTA Amino Acid Sequence

|     |             |             |                 |             |            |
|-----|-------------|-------------|-----------------|-------------|------------|
| 1   | MESPSAPPHR  | WCIPWQRLLL  | TASLLTFWNP      | PTTAKLTIES  | TPFNVAEGKE |
| 51  | VLLLVHNL PQ | HLFGYSWYKG  | ERVDGNRQII      | GYVIGTQQAT  | PGPAYSGREI |
| 101 | IYPNASLLIQ  | NIIQN DTGFY | TLHVIKSDLV      | NEEATGQFRV  | YPELPKPSIS |
| 151 | SNN SKPVEDK | DAVAFTCEPE  | TQDATY LWW      | NNQSLPVSPR  | LQLSNGNRTL |
| 201 | TLFNVTRNDT  | ASYKCETQNP  | VSARRSDSVI      | LNVL YGPDAP | TISPLNTSYR |
| 251 | SGENLNLSCH  | AASNPPAQYS  | W FVNGTFQQS     | TQELFIPNIT  | VNNSGSYTCQ |
| 301 | AHNSDTGLNR  | TTVTTITVYA  | EPPKPFITSN      | NSNPVEDEDA  | VALTCEPEIQ |
| 351 | NTTYLWVVNN  | QSLPVSPRLQ  | LSNDNRTLTL      | LSVTRNDVGP  | YECGIQNELS |
| 401 | VDHSDPVILN  | VLYGPDDPTI  | SPSYTYRPG       | VNLSLSCHAA  | SNPPAQYSWL |
| 451 | IDGNIQQHTQ  | ELFISNITEK  | NSGLYTCQAN      | NSASGHSRTT  | VKTITVSAEL |
| 501 | PKPSISSNNS  | KPVEDKDAVA  | FTCEPEAQNT      | TYLWVWNGQS  | LPVSPRLQLS |
| 551 | NGNRTLTLFN  | VTRNDARAYV  | CGIQNSVSAN      | RSDPVTLDVL  | YGPDTPIISP |
| 601 | PDSSYLSGAN  | LNLSCHSASN  | PSPQYSWRIN      | GIPQQHTQVL  | FIAKITPNNN |
| 651 | GTYACFVSNL  | ATGRNNSIVK  | SITVSASGTL      | VNGDKLYRAD  | SRPPDEIKRS |
| 701 | GGLMPRGHNE  | YFDRGTQMNI  | NLYDHARGTQ      | TGFVRYDDGY  | VSTSLSLRSA |
| 751 | HLAQQSILSG  | YSTYYIYVIA  | TAPNMFNVND      | VLGVYSPHPY  | EQEVSALGGI |
| 801 | PYSQIYGWYR  | VNFGVIDERL  | HRNREYRDRY      | YRNLNIAPAE  | DGYRLAGFPP |
| 851 | DHQAWREEPW  | IHHAPQGCGN  | SSRTITDDTC      | NEETQNLSTI  | YLRKYQSKVK |
| 901 | RQIFSDYQSE  | VDIYNRIRDE  | L (SEQ ID NO:8) |             |            |

FIG.2B

5/62

## hCEA-LTB Coding Sequence

```

1   ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCAAGCTCAC TATTGAATCC ACGCCGTTCA ATGTCGCAGA GGGGAAGGAG
151 GTGCTTCTAC TTGTCCACAA TCTGCCCCAG CATCTTTTGT GCTACAGCTG
201 GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA GGATATGTAA
251 TAGGAACTCA ACAAGCTACC CCAGGGCCCCG CATAAGTGG TCGAGAGATA
301 ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC
351 AGGATTCTAC ACCCTACACG TCATAAAGTC AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTCCATCTCC
451 AGCAACAACCT CCAAACCCGT GGAGGACAAG GATGCTGTGG CCTTCACCTG
501 TGAACCTGAG ACTCAGGACG CAACCTACCT GTGGTGGGTA AACAAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CAGGACCCTC
601 ACTCTATTCA ATGTCACAAG AAATGACACA GCAAGCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGCCA GGCGCAGTGA TTCAGTCATC CTGAATGTCC
701 TCTATGGCCC GGATGCCCCC ACCATTTCCC CTCTAAACAC ATCTTACAGA
751 TCAGGGGAAA ATCTGAACCT CTCCTGCCAC GCAGCCTCTA ACCCACCTGC
801 ACAGTACTCT TGGTTTGTCA ATGGGACTTT CCAGCAATCC ACCCAAGAGC
851 TCTTTATCCC CAACATCACT GTGAATAATA GTGGATCCTA TACGTGCCAA
901 GCCCATAACT CAGACACTGG CCTCAATAGG ACCACAGTCA CGACGATCAC
951 AGTCTATGCA GAGCCACCCA AACCCTTCAT CACCAGCAAC AACTCCAACC
1001 CCGTGGAGGA TGAGGATGCT GTAGCCTTAA CCTGTGAACC TGAGATTGAG
1151 AACACAACCT ACCTGTGGTG GGTAATAAT CAGAGCCTCC CGGTGAGTCC
1101 CAGGCTGCAG CTGTCCAATG ACAACAGGAC CCTCACTCTA CTCAGTGTCA
1151 CAAGGAATGA TGTAGGACCC TATGAGTGTG GAATCCAGAA CGAATTAAGT
1201 GTTGACCACA GCGACCCAGT CATCCTGAAT GTCCTCTATG GCCCAGACGA
1251 CCCCACCATT TCCCCCTCAT ACACCTATTA CCGTCCAGGG GTGAAECTCA
1301 GCCTCTCCTG CCATGCAGCC TCTAACCAC CTGCACAGTA TTCTTGGCTG
1351 ATTGATGGGA ACATCCAGCA ACACACACAA GAGCTCTTTA TCTCGAACAT
1401 CACTGAGAAG AACAGCGGAC TCTATACCTG CCAGGCCAAT AACTCAGCCA
1451 GTGGCCACAG CAGGACTACA GTCAAGACAA TCACAGTCTC TGGGAGCTG
1501 CCCAAGCCCT CCATCTCCAG CAACAACCTCC AAACCCGTGG AGGACAAAGGA
1551 TGCTGTGGCC TTCACCTGTG AACCTGAGGC TCAGAACACA ACCTACCTGT
1601 GGTGGGTAAA TGGTCAGAGC CTCCCAGTCA GTCCCAGGCT GCAGCTGTCC
1651 AATGGCAACA GGACCCTCAC TCTATTCAAT GTCACAAGAA ATGACGCAAG
1701 AGCCTATGTA TGTGGAATCC AGAACTCAGT GAGTGCAAAC CGCAGTGACC
1751 CAGTCACCCT GGATGTCCTC TATGGGCCGG ACACCCCAT CATTTCCTCC
1801 CCAGACTCGT CTTACCTTTC GGGAGCGAAC CTCAACCTCT CCTGCCACTC
1851 GGCCTCTAAC CCATCCCCGC AGTATTCTTG GCGTATCAAT GGGATACCGC

```

FIG.3A-1

6/62

```
1901 AGCAACACAC ACAAGTTCTC TTTATCGCCA AAATCACGCC AAATAATAAC
1951 GGGACCTATG CCTGTTTTGT CTCTAACTTG GCTACTGGCC GCAATAATTC
2001 CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACTCTA GATGCTCCCC
2051 AGTCTATTAC AGAACTATGT TCGGAATATC GCAACACACA AATATATACG
2101 ATAAATGACA AGATACTATC ATATACGGAA TCGATGGCAG GTAAAAGAGA
2151 AATGGTTATC ATTACATTTA AGAGCGGCGC AACATTTTCAG GTCGAAGTCC
2201 CGGGCAGTCA ACATATAGAC TCCCAAAAAA AAGCCATTGA AAGGATGAAG
2251 GACACATTAA GAATCACATA TCTGACCGAG ACCAAAATTG ATAAATTATG
2301 TGTATGGAAT AATAAAACCC CCAATTCAAT TGC GGCAATC AGTATGGAAA
      ACTAG (SEQ ID NO:9)
```

FIG.3A-2

7/62

## CEA-LTB Amino Acid Sequence

|     |             |             |            |                     |            |
|-----|-------------|-------------|------------|---------------------|------------|
| 1   | MESPSAPPHR  | WCIPWORLLL  | TASLLTFWNP | PTTAKLTIES          | TPFNVAEGKE |
| 51  | VLLLVHNL PQ | HLFGYSWYKG  | ERVDGNRQII | GYVIGTQQAT          | PGPAYSGREI |
| 101 | IYPNASLLIQ  | NIIQNDTG FY | TLHVIKSDLV | NEEATGQFRV          | YPELPKPSIS |
| 151 | SNNSKPVEDK  | DAVAFTCEPE  | TQDATYLWWV | NNQSLPVSPR          | LQLSNGNRTL |
| 201 | TLFNVTRNDT  | ASYKCETQNP  | VSARRSDSVI | LNVL YGPDAP         | TISPLNTSYR |
| 251 | SGENLNLSCH  | AASNPPAQYS  | WFVNGTFQQS | TQELFIPNIT          | VNNSGSYTCQ |
| 301 | AHNSDTGLNR  | TTVTTITVYA  | EPPKPFITSN | NSNPVEDEDA          | VALTCEPEIQ |
| 351 | NTTYLWWVNN  | QSLPVSPRLQ  | LSNDNRTLTL | LSVTRNDVGP          | YECGIQNELS |
| 401 | VDHSDPVILN  | VLYGPDDPTI  | SPSYTYRPG  | VNLSLSCHAA          | SNPPAQYSWL |
| 451 | IDGNIQQHTQ  | ELFISNITEK  | NSGLYTCQAN | NSASGHSRTT          | VKTITVSAEL |
| 501 | PKPSISSNNS  | KPVEDKDAVA  | FTCEPEAQNT | TYLWWVNGQS          | LPVSPRLQLS |
| 551 | NGNRTLTLFN  | VTRNDARAYV  | CGIQNSVSAN | RSDPVTLDVL          | YGPDTPIISP |
| 601 | PDSSYLSGAN  | LNLSCHSASN  | PSPQYSWRIN | GIPQQHTQVL          | FIAKITPNNN |
| 651 | GTYACFVSNL  | ATGRNNSIVK  | SITVSASGTL | DAPQSITELC          | SEYRNTQIYT |
| 701 | INDKILSYTE  | SMAGKREMI   | ITFKSGATFQ | VEVPGSQHID          | SQKKAIERMK |
| 751 | DTLRITYLTE  | TKIDKLCVWN  | NKTPNSIAAI | SMEN (SEQ ID NO:10) |            |

FIG.3B

8/62

## CEAopt-LTB Nucleotide Sequence

```

1   ATGGAGAGCC CCAGCGCCCC CCCCCACCGC TGGTGCATCC CCTGGCAGCG
   CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG
101  CCAAGCTGAC CATCGAGAGC ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG
   GTGCTGCTGC TGGTGCACAA CCTGCCCCAG CACCTGTTCG GCTACAGCTG
201  GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC GGCTACGTGA
   TCGGCACCCA GCAGGCCACC CCCGGCCCCG CCTACAGCGG CCGCGAGATC
301  ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC
   CGGCTTCTAC ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG
401  CCACCGGCCA GTTCCGCGTG TACCCGAGC TGCCCAAGCC CAGCATCAGC
   AGCAACAACA GCAAGCCCGT GGAGGACAAG GACGCCGTGG CCTTCACCTG
501  CGAGCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG AACAACCAGA
   GCCTGCCCGT GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCCCTG
601  ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC
   CCAGAACCCC GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC
701  TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CAGCTACCGC
   AGCGGCGAGA ACCTGAACCT GAGCTGCCAC GCCGCCAGCA ACCCCCCCGC
801  CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC
   TGTTTCATCC CAACATCACC GTGAACAACA GCGGCAGCTA CACCTGCCAG
901  GCCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC
   CGTGTACGCC GAGCCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC
1001 CCGTGGAGGA CGAGGACGCC GTGGCCCTGA CCTGCGAGCC CGAGATCCAG
   AACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGC CCGTGAGCCC
1101 CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG CTGAGCGTGA
   CCCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC
1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCCGACGA
   CCCCACCATC AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA
1301 GCCTGAGCTG CCACGCCGCC AGCAACCCCC CCGCCCAGTA CAGCTGGCTG
   ATCGACGGCA ACATCCAGCA GCACACCCAG GAGCTGTTCA TCAGCAACAT
1401 CACCGAGAAG AACAGCGGCC TGTACACCTG CCAGGCCAAC AACAGCGCCA
   GCGGCCACAG CCGCACCAAC GTGAAGACCA TCACCGTGAG CGCCGAGCTG
1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA
   CGCCGTGGCC TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT
1601 GGTGGGTGAA CGGCCAGAGC CTGCCCGTGA GCCCCCGCCT GCAGCTGAGC
   AACGGCAACC GCACCCTGAC CCTGTTCAAC GTGACCCGCA ACGACGCCCG
1701 CGCCTACGTG TGCGGCATCC AGAACAGCGT GAGCGCCAAC CGCAGCGACC
   CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCAT CATCAGCCCC
1801 CCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG
   CGCCAGCAAC CCCAGCCCCC AGTACAGCTG GCGCATCAAC GGCATCCCCC

```

FIG.4A



9/62

```
1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC
      GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAGCTCCCC
      AGACTATTAC AGAACTATGT TCGGAATATC GCAACACACA AATATATACG
2101 ATAAATGACA AGATACTATC ATATACGGAA TCGATGGCAG GCAAAAGAGA
      AATGGTTATC ATTACATTTA AGAGCGGCGA AACATTTTCAG GTCGAAGTCC
2201 CGGGCAGTCA ACATATAGAC TCCAGAAAAA AAGCCATTGA AAGGATGAAG
      GACACATTAA GAATCACATA TCTGACCGAG ACCAAAATTG ATAAATTATG
2301 TGTATGGAAT AATAAAACCC CCAATTCAAT TGCGGCAATC AGTATGGAAA
      ACTAG (SEQ ID NO:11)
```

FIG.4B

10/62

## hCEA-LTBopt Coding Sequence

```

1   ATGGAGAGCC CCAGCGCCCC CCCCCACCGC TGGTGCATCC CCTGGCAGCG
   CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG
101 CCAAGCTGAC CATCGAGAGC ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG
   GTGCTGCTGC TGGTGCACAA CCTGCCCCAG CACCTGTTCG GCTACAGCTG
201 GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC GGCTACGTGA
   TCGGCACCCA GCAGGCCACC CCGGGCCCCG CCTACAGCGG CCGCGAGATC
301 ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC
   CGGCTTCTAC ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG
401 CCACCGGCCA GTTCCGCGTG TACCCCGAGC TGCCCAAGCC CAGCATCAGC
   AGCAACAACA GCAAGCCCGT GGAGGACAAG GACGCCGTGG CCTTCACCTG
501 CGAGCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG AACAACCAGA
   GCCTGCCCCG GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCCCTG
601 ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC
   CCAGAACCCC GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC
701 TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CAGCTACCGC
   AGCGGCGAGA ACCTGAACCT GAGCTGCCAC GCCGCCAGCA ACCCCCCCGC
801 CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC
   TGTTTCATCCC CAACATCACC GTGAACAACA GCGGCAGCTA CACCTGCCAG
901 GCCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC
   CGTGTACGCC GAGCCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC
1001 CCGTGGAGGA CGAGGACGCC GTGGCCCTGA CCTGCGAGCC CGAGATCCAG
   AACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGC CCGTGAGCCC
1101 CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG CTGAGCGTGA
   CCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC
1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCGACGA
   CCCCACCATC AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA
1301 GCCTGAGCTG CCACGCCGCC AGCAACCCCC CCGCCCAGTA CAGCTGGCTG
   ATCGACGGCA ACATCCAGCA GCACACCCAG GAGCTGTTCA TCAGCAACAT
1401 CACCGAGAAG AACAGCGGCC TGTAACCTG CCAGGCCAAC AACAGCGCCA
   GCGGCCACAG CCGCACACC GTGAAGAACA TCACCGTGAG CGCCGAGCTG
1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA
   CGCCGTGGCC TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT
1601 GGTGGGTGAA CGGCCAGAGC CTGCCCGTGA GCCCCCGCCT GCAGCTGAGC
   AACGGCAACC GCACCCTGAC CCTGTTCAAC GTGACCCGCA ACGACGCCCC
1701 CGCCTACGTG TCGGGCATCC AGAACAGCGT GAGCGCCAAC CGCAGCGACC
   CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCAT CATCAGCCCC
1801 CCCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG
   CGCCAGCAAC CCCAGCCCCC AGTACAGCTG GCGCATCAAC GGCATCCCCC

```

FIG.5A-1

11/62

1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC  
GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG  
2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAGCCCCC  
AGAGCATCAC CGAGCTGTGC AGCGAGTACC GGAACACCCA GATCTACACC  
2101 ATCAACGACA AGATCCTGAG CTACACCGAG AGCATGGCCG GCAAGAGGGA  
GATGGTGATC ATCACCTTCA AGAGCGGCGC CACCTTCCAG GTGGAGGTGC  
2201 CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA GCGGATGAAG  
GACACCCTGC GGATCACCTA CCTACCGAG ACCAAGATCG ACAAGCTGTG  
2301 CGTGTGGAAC AACAAGACCC CCAACAGCAT CGCCGCCATC AGCATGGAGA  
ATTGATAA (SEQ ID NO:12)

FIG.5A-2

12/62

## hCEA-LTB Amino Acid Sequence

|     |             |            |            |                     |            |
|-----|-------------|------------|------------|---------------------|------------|
| 1   | MESPSAPPHR  | WCIPWQRLLL | TASLLTFWNP | PTTAKLTIES          | TPFNVAEGKE |
| 51  | VLLL VHNLPQ | HLFGYSWYKG | ERVDGNRQII | GYVIGTQQAT          | PGPAYSGREI |
| 101 | IYPNASLLIQ  | NIIQNDTGFY | TLHVIKSDLV | NEEATGQFRV          | YPELPKPSIS |
| 151 | SNNSKPVEDK  | DAVAFTCEPE | TQDATYLWWV | NNQSLPVSPR          | LQLSNGNRTL |
| 201 | TLFNVTRNDT  | ASYKCETQNP | VSARRSDSVI | LVVLYGPDAP          | TISPLNTSYR |
| 251 | SGENLNLSCH  | AASNPPAQYS | WVNGTFQQS  | TQELFIPNI           | VNNSGSYTCQ |
| 301 | AHNSDTGLNR  | TTVTTITVYA | EPPKPFITSN | NSNPVEDEDA          | VALTCEPEIQ |
| 351 | NTTYLWWVNN  | QSLPVSPRLQ | LSNDNRTLTL | LSVTRNDVGP          | YECGIQNELS |
| 401 | VDHSDPVILN  | VLYGPDDPTI | SPSYTYRPG  | VNLSLSCHAA          | SNPPAQYSWL |
| 451 | IDGNIQQHTQ  | ELFISNITEK | NSGLYTCQAN | NSASGHSRTT          | VKTITVSAEL |
| 501 | PKPSISSNNS  | KPVEDKDAVA | FTCEPEAQNT | TYLWWVNGQS          | LPVSPRLQLS |
| 551 | NGNRTLTLFN  | VTRNDARAYV | CGIQNSVSAN | RSDPVTLDVL          | YGPDTPIISP |
| 601 | PDSSYLSGAN  | LNLSCHSASN | PSPQYSWRIN | GIPQQHTQVL          | FIAKITPNNN |
| 651 | GTYACFVSNL  | ATGRNNSIVK | SITVSASGTS | RA PQSITE LC        | SEYRNTQIYT |
| 701 | INDKILSYTE  | SMAGKREMI  | ITFKSGATFQ | VEVPGSQHID          | SQKKAIERMK |
| 751 | DTLRITYLTE  | TKIDKLCVWN | NKTPNSIAAI | SMEN (SEQ ID NO:13) |            |

FIG.5B

13/62

## Rhesus CEAopt-LTBopt Coding Sequence

ATGGGCAGCC CCAGCGCCCC CCTGCACCGC TGGTGCATCC CCTGGCAGAC CCTGCTGCTG ACGGCCAGCC  
 TGCTGACCTT CTGGAACCCC CCCACCACCG CCCAGCTGAC CATCGAGAGC CGCCCCCTTCA ACGTGGCCGA  
 GGGCAAGGAG GTGCTGCTGC TGGCCCCACAA CGTGAGCCAG AACCTGTTCC GCTACATCTG GTACAAGGGC  
 GAGCGGTGG ACGCCAGCCG CCGCATCGGC AGCTGCGTGA TCCGCACCCA GCAGATCACC CCGGGCCCCG  
 CCCACAGGG CCGCGAGACC ATCGACTTCA ACGCCAGCCT GCTGATCCAC AACGTGACCC AGAGCGACAC  
 CGGCAGCTAC ACCATCCAGG TGATCAAGGA GGACCTGGTG AACGAGGAGG CCACCGGCCA GTTCCGCGTG  
 TACCCCGAGC TGCCCAAGCC CTACATCAGC AGCAACAACA GCAACCCCGT GGAGGACAAG GACGCCGTGG  
 CCCTGACCTG CGAGCCCGAG ACCCAGGACA CCACCTACCT GTGGTGGTG AACAAACCAGA GCCTGCCCGT  
 GAGCCCCCGC CTGGAGCTGA GCAGCGACAA CCGCACCCCTG ACCGTGTTCA ACATCCCCCG CAACGACACC  
 ACCAGCTACA AGTGGGAGAC CCAGAACCCC GTGAGCGTGC GCCGCAGGA CCCCCTGACC CTGAACGTGC  
 TGTACGGCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CCCCTACCGC GCCGGCGAGA ACCTGAACCT  
 GACCTGCCAC GCCGCCAGCA ACCCCACCGC CCAGTACTTC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC  
 ACCCAGGAGC TGTTCAATCC CAACATCACC GTGAACAACA GCGGCAGCTA CATGTGCCAG GCCCACAACA  
 GCGCCACCGG CCTGAACCGC ACCACCGTGA CCGCCATCAC CGGTACGCC GAGCTGCCCA AGCCCTACAT  
 CACCAGCAAC AACAGCAACC CCATCGAGGA CAAGGACGCC GTGACCCCTGA CCTGCGAGCC CGAGACCCAG  
 GACACCACT ACCTGTGGTG GGTGAACAAC CAGAGCCTGA GCGTGAGCAG CCGCCTGGAG CTGAGCAACG  
 ACAACCGCAC CTGACCGTG TTCAACATCC CCGCAACGA CACCACCTTC TACGAGTCCG AGACCCAGAA  
 CCCCCTGAGC GTGGCGCGCA GCGACCCCGT GACCCTGAAC GTGCTGTAGG GCCCCGACGC CCCCACCATC  
 AGCCCCCTGA ACACCCCCTA CCGCGCCGGC GAGAACCTGA ACCTGAGCTG CCACGCCGCC AGCAACCCCG  
 CCGCCAGTA CAGCTGGTTC GTGAACGGCA CCTTCCAGCA GAGCACCCAG GAGCTGTCA TCCCCAACAT  
 CACCGTGAAC AACAGCGGCA GCTACATGTG CCAGGCCAC AACAGCGCA CCGCCTGAA CCGCACCCAC  
 GTGACCGCCA TCACCGTGA CGTGGAGCTG CCCAAGCCCT ACATCAGCAG CAACAACAGC AACCCCATCG

FIG.6A-1

AGGACAAGGA CGCCGTGACC CTGACCTGGC AGCCCGTGGC CGAGAACACC ACCTACCTGT GGTGGTGAA  
CAACCAGAGC CTGAGCGTGA GCCCCCGCCT GCAGCTGAGC AACGGCAACC GCATCCTGAC CCTGCTGAGC  
GTGACCCGCA ACGACACCGG CCCCTACGAG TCGGGCATCC AGAACAGCGA GAGCGCCAAG CGCAGCGACC  
CCGTGACCCT GAACGTGACC TACGGCCCCG ACACCCCCAT CATCAGCCCC CCCGACCTGA GCTACCGCAG  
CGGCGCCAAC CTGAACCTGA GCTGCCACAG CGACAGCAAC CCCAGCCCCC AGTACAGCTG GCTGATCAAC  
GGCACCCCTGC GCCAGCACAC CCAGGTGCTG TTCATCAGCA AGATCACCAG CAACAACAGC GGCGCCCTACG  
CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG CATCGTGAAG AACATCAGCG TGAGCAGCGG  
CGACAGCTCT AGAGCCCCC AGAGCATCAC CGAGCTGTGC AGCGAGTACC GGAACACCCA GATCTACACC  
ATCAACGACA AGATCCTGAG CTACACCGAG AGCATGGCCG GCAAGAGGGA GATGGTGATC ATCACCTTCA  
AGAGCGGCGC CACCTTCCAG GTGGAGGTGC CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA  
GCGGATGAAG GACACCCTGC GGATCACCTA CCTACCCGAG ACCAAGATCG ACAAGCTGTG CGTGTGGAAC  
ACAAGACCC CCAACAGCAT CGCCGCCATC AGCATGGAGA ATTGATAA (SEQ ID NO:14)

FIG. 6A-2

15/62

## RhCEAopt-LTBopt Amino Acid Sequence

```
1   MGSPSAPLHR WCIPWQTLLE TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWYKG ERVDASRRIG SCVIRTOQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELPPKPYIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFVNGTFQQS TQELFIPNIT VNNSGSYMCQ
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQQSTQ ELFIPNITVN NSGSYMCQAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPEY CGIQNSSESAK RSDPVTNLNV YGPDTPIIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSS RAPQSITELC SEYRNTQIYT
701 INDKILSYTE SMAGKREIVI ITFKSGATFQ VEVPGSQHID SQKKAIERMK
751 DTLRITYLTE TKIDKLCVWN NKTPNSIAAI SMEN (SEQ ID NO:15)
```

FIG.6B

16/62

## Nucleotide Sequence of First Rhesus Monkey CEA

```

1   ATGGGGTCTC CCTCAGCCCC TCTTCACAGA TGGTGCATCC CCTGGCAGAC
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCCAGCTCAC TATTGAATCC AGGCCGTTCA ATGTTGCAGA GGGGAAGGAG
151 GTTCTTCTAC TTGCCCACAA TGTGTCCCAG AATCTTTTGT GCTACATTTG
201 GTACAAGGGA GAAAGAGTGG ATGCCAGCCG TCGAATTGGA TCATGTGTAA
251 TAAGAACTCA ACAAATTACC CCAGGGCCCC CACACAGCGG TCGAGAGACA
301 ATAGACTTCA ATGCATCCCT GCTGATCCAC AATGTCACCC AGAGTGACAC
351 AGGATCCTAC ACCATACAAG TCATAAAGGA AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTACATCTCC
451 AGCAACAAC TCCAACCCGT GGAGGACAAG GATGCTGTGG CCTTAACCTG
501 TGAACCTGAG ACTCAGGACA CAACCTACCT GTGGTGGGTA AACAAACAGA
551 GCCTCCCGGT CAGTCCCAGG CTGGAGCTGT CCAGTGACAA CAGGACCCCTC
601 ACTGTATTCA ATATTCCAAG AAATGACACA ACATCCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGTCA GACGCAGCGA CCCAGTCACC CTGAACGTCC
701 TCTATGGCCC GGATGCGCCC ACCATTTCCC CTCTAAACAC ACCTTACAGA
751 GCAGGGGAAA ATCTGAACCT CACCTGCCAC GCAGCCTCTA ACCCAACTGC
801 ACAGTACTTT TGGTTTGTCA ATGGGACGTT CCAGCAATCC ACACAAGAGC
851 TCTTTATACC CAACATCACC GTGAATAATA GCGGATCCTA TATGTGCCAA
901 GCCCATAACT CAGCCACTGG CCTCAATAGG ACCACAGTCA CGGCGATCAC
951 AGTCTACGCG GAGCTGCCCA AGCCCTACAT CACCAGCAAC AACTCCAACC
1001 CCATAGAGGA CAAGGATGCT GTGACCTTAA CCTGTGAACC TGAGACTCAG
1051 GACACAACCT ACCTGTGGTG GGTAAACAAT CAGAGCCTCT CGGTCAGTTC
1101 CAGGCTGGAG CTGTCCAATG ACAACAGGAC CCTCACTGTA TTCAATATTC
1151 CAAGAAACGA CACAACGTTC TACGAATGTG AGACCCAGAA CCCAGTGAGT
1201 GTCAGACGCA GCGACCCAGT CACCCTGAAT GTCCTCTATG GCCCGGATGC
1251 GCCCACCATT TCCCCTCTAA ACACACCTTA CAGAGCAGGG GAAAATCTGA
1301 ACCTCTCCTG CCACGCAGCC TCTAACCCAG CTGCACAGTA CTCTTGGTTT
1351 GTCAATGGGA CGTTCCAGCA ATCCACACAA GAGCTCTTTA TAACCAACAT
1401 CACCGTGAAT AATAGCGGAT CCTATATGTG CCAAGCCCAT AAETCAGCCA
1451 CTGGCCTCAA TAGGACCACA GTCACGGCGA TCACAGTCTA TGTGGAGCTG
1501 CCCAAGCCCT ACATCTCCAG CAACAACCTC AACCCCATAG AGGACAAGGA
1551 TGCTGTGACC TTAACCTGTG AACCTGTGGC TGAGAACACA ACCTACCTGT
1601 GGTGGGTAAC CAATCAGAGC CTCTCGGTCA GTCCCAGGCT GCAGCTCTCC
1651 AATGGCAACA GGATCCTCAC TCTACTCAGT GTCACACGGA ATGACACAGG
1701 ACCCTATGAA TGTGGAATCC AGAACTCAGA GAGTGCAAAA CGCAGTGACC
1751 CAGTCACCCT GAATGTCACC TATGGCCCAG ACACCCCAT CATATCCCCC
1801 CCAGACTTGT CTTACCGTTC GGGAGCAAAC CTCAACCTCT CCTGCCACTC

```

FIG.7A-1



17/62

```
1851  GGA CTCTAAC CCATCCCCGC AGTATTCTTG GCTTATCAAT GGGACACTGC
1901  GGCAACACAC ACAAGTTCTC TTTATCTCCA AAATCACATC AAACAATAGC
1951  GGGGCCTATG CCTGTTTTGT CTCTAACTTG GCTACCGGTC GCAATAACTC
2001  CATAGTCAAG AACATCTCAG TCTCCTCTGG CGATTCAGCA CCTGGAAGTT
2051  CTGGTCTCTC AGCTAGGGCT ACTGTCGGCA TCATAATTGG AATGCTGGTT
2101  GGGGTTGCTC TGATGTAG (SEQ ID NO:16)
```

FIG.7A-2

18/62

## Nucleotide Sequence of Second Rhesus Monkey CEA

```

1   ATGGGGTCTC CCTCAGCCCC TCTTCACAGA TGGTGCATCC CCTGGCAGAC
51  GCTCCTGCTC ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG
101 CCCAGCTCAC TATTGAATCC AGGCCGTTCA ATGTTGCAGA GGGGAAGGAG
151 GTTCTTCTAC TTGCCCACAA TGTGTCCCAG AATCTTTTTG GCTACATTTG
201 GTACAAGGGA GAAAGAGTGG ATGCCAGCCG TCGAATTGGA TCATGTGTAA
251 TAAGAACTCA ACAAATTACC CCAGGGCCCG CACACAGCGG TCGAGAGACA
301 ATAGACTTCA ATGCATCCCT GCTGATCCAC AATGTCACCC AGAGTGACAC
351 AGGATCCTAC ACCATACAAG TCATAAAGGA AGATCTTGTG AATGAAGAAG
401 CAACTGGCCA GTTCCGGGTA TACCCGGAGC TGCCCAAGCC CTACATCTCC
451 AGCAACAACCT CCAACCCCGT GGAGGACAAG GATGCTGTGG CCTTAACCTG
501 TGAACCTGAG ACTCAGGACA CAACCTACCT GTGGTGGGTA AACAATCAGA
551 GCCTCCCGGT CAGTCCCAGG CTGGAGCTGT CCAGTGACAA CAGGACCCCTC
601 ACTGTATTCA ATATTCCAAG AAATGACACA ACATCCTACA AATGTGAAAC
651 CCAGAACCCA GTGAGTGTCA GACGCAGCGA CCCAGTCACC CTGAACGTCC
701 TCTATGGCCC GGATGCGCCC ACCATTTCCC CTCTAAACAC ACCTTACAGA
751 GCAGGGGAAA ATCTGAACCT CACCTGCCAC GCAGCCTCTA ACCCAACTGC
801 ACAGTACTTT TGGTTTGTCA ATGGGACGTT CCAGCAATCC ACACAAGAGC
851 TCTTTATACC CAACATCACC GTGAATAATA GCGGATCCTA TATGTGCCAA
901 GCCCATAACT CAGCCACTGG CCTCAATAGG ACCACAGTCA CGGCGATCAC
951 AGTCTACGCG GAGCTGCCCA AGCCCTACAT CACCAGCAAC AACTCCAACC
1001 CCATAGAGGA CAAGGATGCT GTGACCTTAA CCTGTGAACC TGAGACTCAG
1051 GACACAACCT ACCTGTGGTG GGTAAACAAT CAGAGCCTCT CGGTCAGTTC
1101 CAGGCTGGAG CTGTCCAATG ACAACAGGAC CCTCACTGTA TTCAATATTC
1151 CAAGAAACGA CACAACGTTC TACGAATGTG AGACCCAGAA CCCAGTGAGT
1201 GTCAGACGCA GCGACCCAGT CACCCTGAAT GTCTCTATG GCCCGGATGC
1251 GCCCACCATT TCCCCTCTAA ACACACCTTA CAGAGCAGGG GAAAATCTGA
1301 ACCTCTCCTG CCACGCAGCC TCTAACCCAG CTGCACAGTA CTTTGTGTTT
1351 GTCAATGGGA CGTTCCAGCA ATCCACACAA GAGCTCTTTA TACCCAACAT
1401 CACCGTGAAT AATAGCGGAT CCTATATGTG CCAAGCCCAT AACTCAGCCA
1451 CTGGCCTCAA TAGGACCACA GTCACGGCGA TCACAGTCTA TGTGGAGCTG
1501 CCAAGCCCT ACATCTCCAG CAACAACCTC AACCCCATAG AGGACAAGGA
1551 TGCTGTGACC TTAACCTGTG AACCTGTGGC TGAGAACACA ACCTACCTGT
1601 GGTGGGTAAA CAATCAGAGC CTCTCGGTCA GTCCAGGCT GCAGCTCTCC
1651 AATGGCAACA GGATCCTCAC TCTACTCAGT GTCACACGGA ATGACACAGG
1701 ACCCTATGAA TGTGGAATCC AGAACTCAGA GAGTGCAAAA CGCAGTGACC
1751 CAGTCACCCT GAATGTCACC TATGGCCCAG ACACCCCAT CATATCCCCC
1801 CCAGACTTGT CTTACCGTTC GGGAGCAAAC CTCAACCTCT CCTGCCACTC

```

FIG.7B-1

19/62

```
1851  GGACTCTAAC CCATCCCCGC AGTATTCTTG GCTTATCAAT GGGACACTGC
1901  GGCAACACAC ACAAGTTCTC TTTATCTCCA AAATCACATC AAACAATAAC
1951  GGGGCCTATG CCTGTTTTGT CTCTAACTTG GCTACCGGTC GCAATAACTC
2001  CATAGTCAAG AACATCTCAG TCTCCTCTGG CGATTCAGCA CCTGGAAGTT
2051  CTGGTCTCTC AGCTAGGGCT ACTGTCGGCA TCATAATTGG AATGCTGGTT
2101  GGGGTTGCTC TGATGTAG (SEQ ID NO:17)
```

FIG.7B-2

20/62

## Amino Acid Sequence of First Rhesus Monkey CEA Protein

```

1    MGSPSAPLHR WCIPWQTL LLL TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51   VLLLAHNVSQ NLFgyiwykg ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101  IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPEL PKPYIS
151  SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201  TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251  AGENLNLTCH AASNPTAQYF WFVNGTFQQS TQELFIPNIT VNNSGSYMCQ
301  AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351  DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401  VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451  VNGTFQQSTQ ELFIPNITVN NSGSYMCQAH NSATGLNRTT VTAITVYVEL
501  PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551  NGNRILTLLS VTRNDTGPYE CGIQNSESAK RSDPVT LNVT YGPDTPII SP
601  PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651  GAYACFVSNL ATGRNNSIVK NISVSSGDSA PGSSGLSARA TVGIIIGMLV
701  GVALM (SEQ ID NO:18)

```

FIG.7C

## Amino Acid Sequence of Second Rhesus Monkey CEA Protein

```

1    MGSPSAPLHR WCIPWQTL LLL TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51   VLLLAHNVSQ NLFgyiwykg ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101  IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPEL PKPYIS
151  SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201  TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251  AGENLNLTCH AASNPTAQYF WFVNGTFQQS TQELFIPNIT VNNSGSYMCQ
301  AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351  DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401  VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYFWF
451  VNGTFQQSTQ ELFIPNITVN NSGSYMCQAH NSATGLNRTT VTAITVYVEL
501  PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551  NGNRILTLLS VTRNDTGPYE CGIQNSESAK RSDPVT LNVT YGPDTPII SP
601  PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNN
651  GAYACFVSNL ATGRNNSIVK NISVSSGDSA PGSSGLSARA TVGIIIGMLV
701  GVALM (SEQ ID NO:19)

```

FIG.7D

21/62

## Amino Acid Sequence of Human CEA Protein

|     |                   |            |            |            |            |
|-----|-------------------|------------|------------|------------|------------|
| 1   | MESPSAPPHR        | WCIPWQRLLL | TASLLTFWNP | PTTAKLTIES | TPFNVAEGKE |
| 51  | VLLL VHNLPQ       | HLFGYSWKYK | ERVDGNRQII | GYVIGTQQAT | PGPAYSGREI |
| 101 | IYPNASLLIQ        | NIIQNDTGFI | TLHVIKSDLV | NEEATGQFRV | YPELPKPSIS |
| 151 | SNNSKPVEDK        | DAVAFTCEPE | TQDATYLWWV | NNQSLPVSPR | LQLSNGNRTL |
| 201 | TLFNVTRNDT        | ASYKCETQNP | VSARRSDSVI | LVNLYGPDAP | TISPLNTSYR |
| 251 | SGENLNLSCH        | AASNPPAQYS | WVFNQTFQQS | TQELFIPNIT | VNNSGSYTCQ |
| 301 | AHNSDTGLNR        | TTVTTITVYA | EPPKPFITSN | NSNPVEDEDA | VALTCEPEIQ |
| 351 | NTTYLWWVNN        | QSLPVSPRLQ | LSNDNRTLTL | LSVTRNDVGP | YECGIQNELS |
| 401 | VDHSDPVILN        | VLYGPDDPTI | SPSYTYRPG  | VNLSLSCHAA | SNPPAQYSWL |
| 451 | IDGNIQQHTQ        | ELFISNITEK | NSGLYTCQAN | NSASGHSRTT | VKTITVSAEL |
| 501 | PKPSISSNNS        | KPVEDKDAVA | FTCEPEAQNT | TYLWWVNGQS | LPVSPRLQLS |
| 551 | NGNRTLTLFN        | VTRNDARAYV | CGIQNSVSAN | RSDPVTLDVL | YGPDTPIISP |
| 601 | PDSSYLSGAN        | LNLSCHSASN | PSPQYSWRIN | GIPQHTQVL  | FIAKITPNNN |
| 651 | GTYACFVSNL        | ATGRNNSIVK | SITVSASGTS | PGLSAGATVG | IMIGULVGVA |
| 701 | LI (SEQ ID NO:20) |            |            |            |            |

FIG.7E

22/62

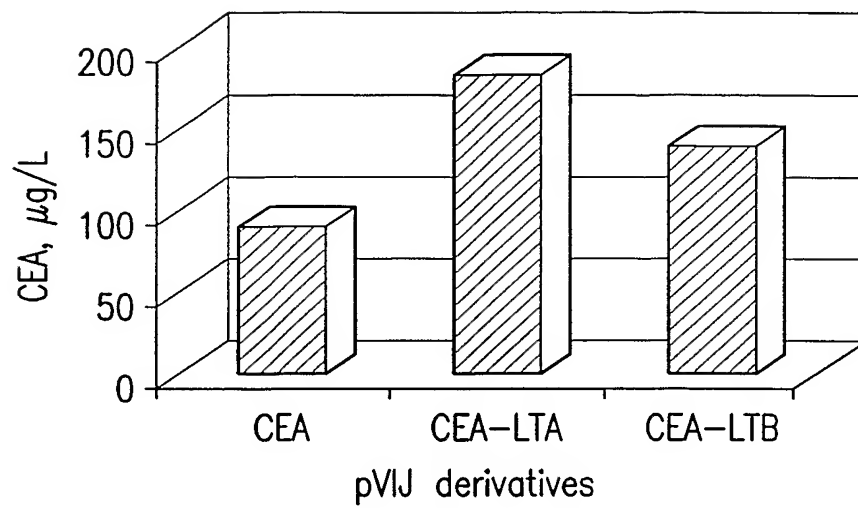


FIG.8A

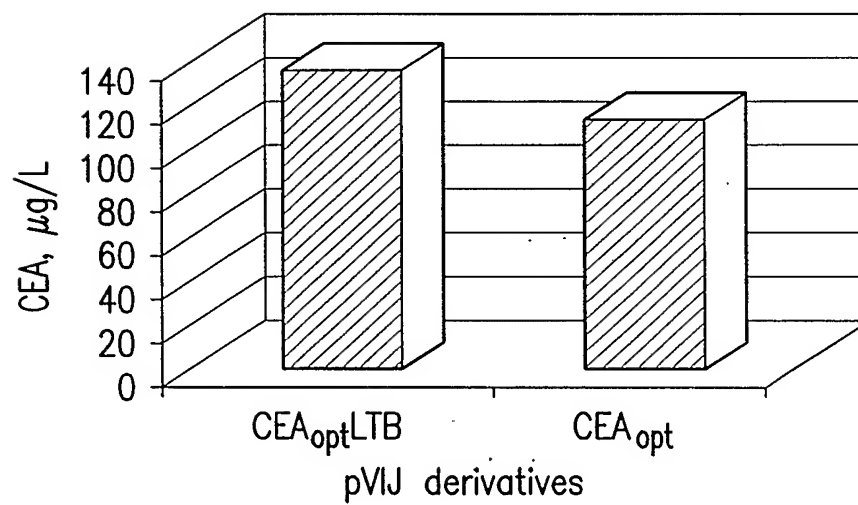
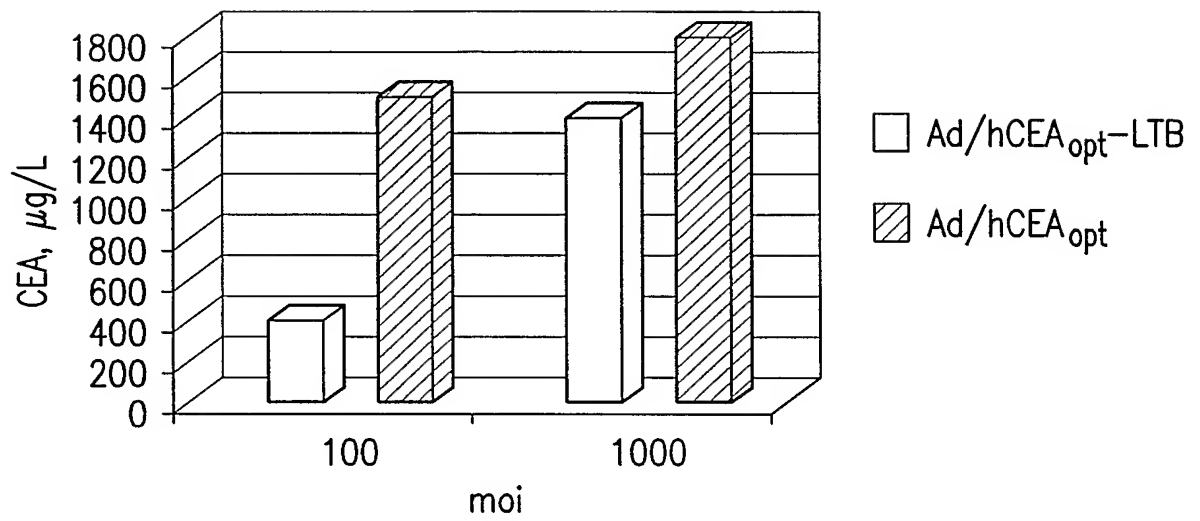


FIG.8B

23/62



Ad vectors  
HeLa cells

FIG.9

24/62

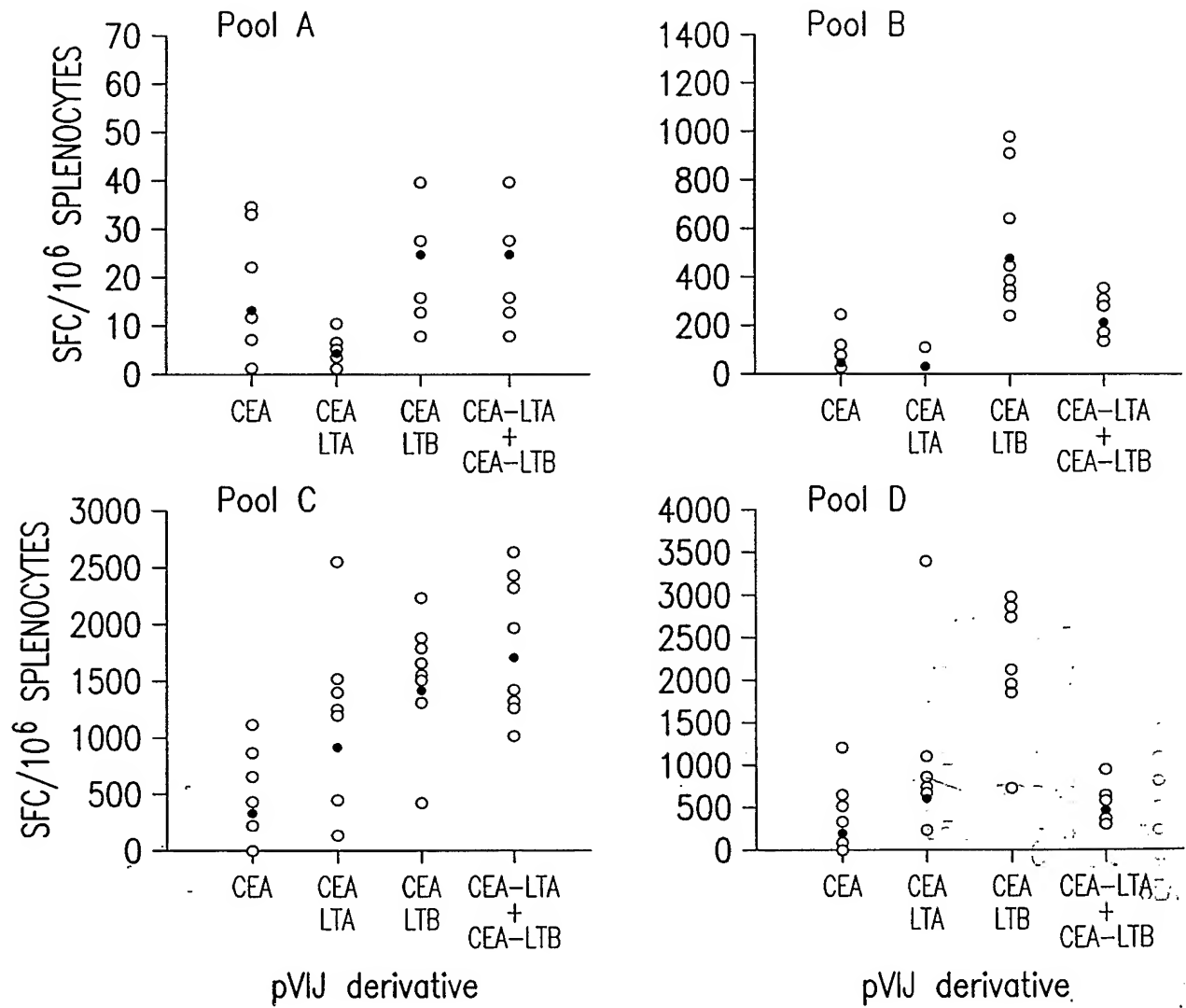


FIG. 10A



25/62

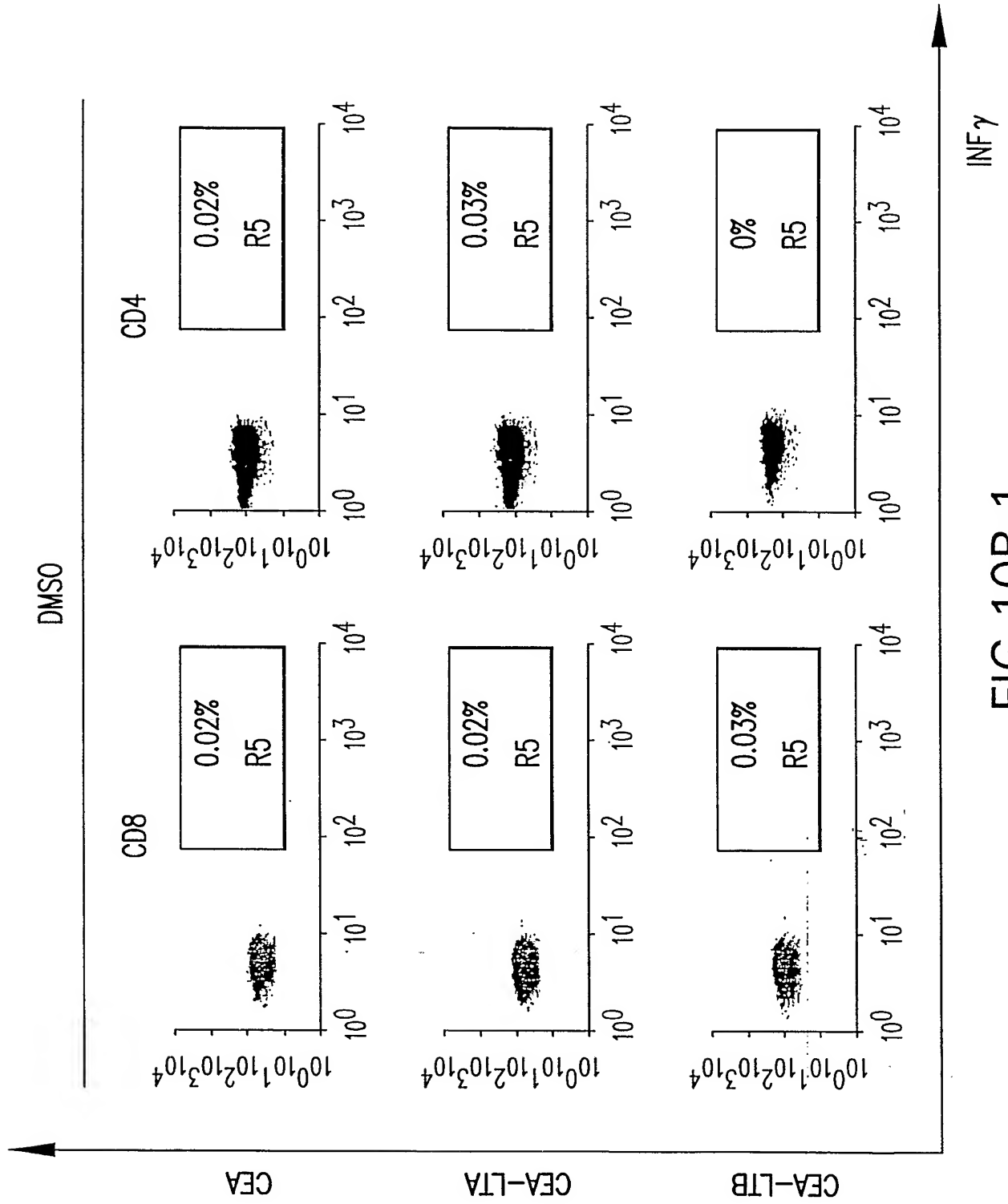
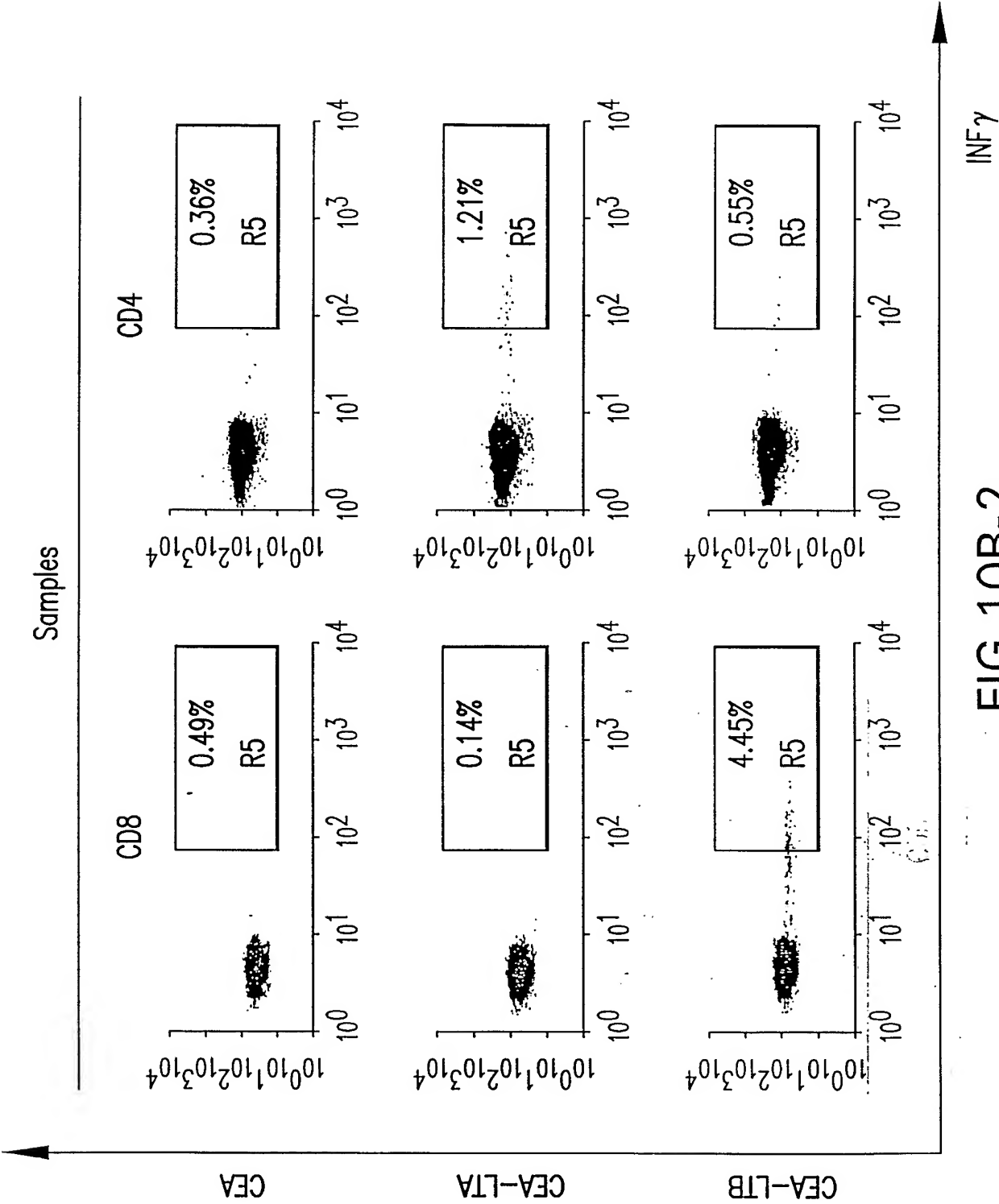


FIG.10B-1



27/62

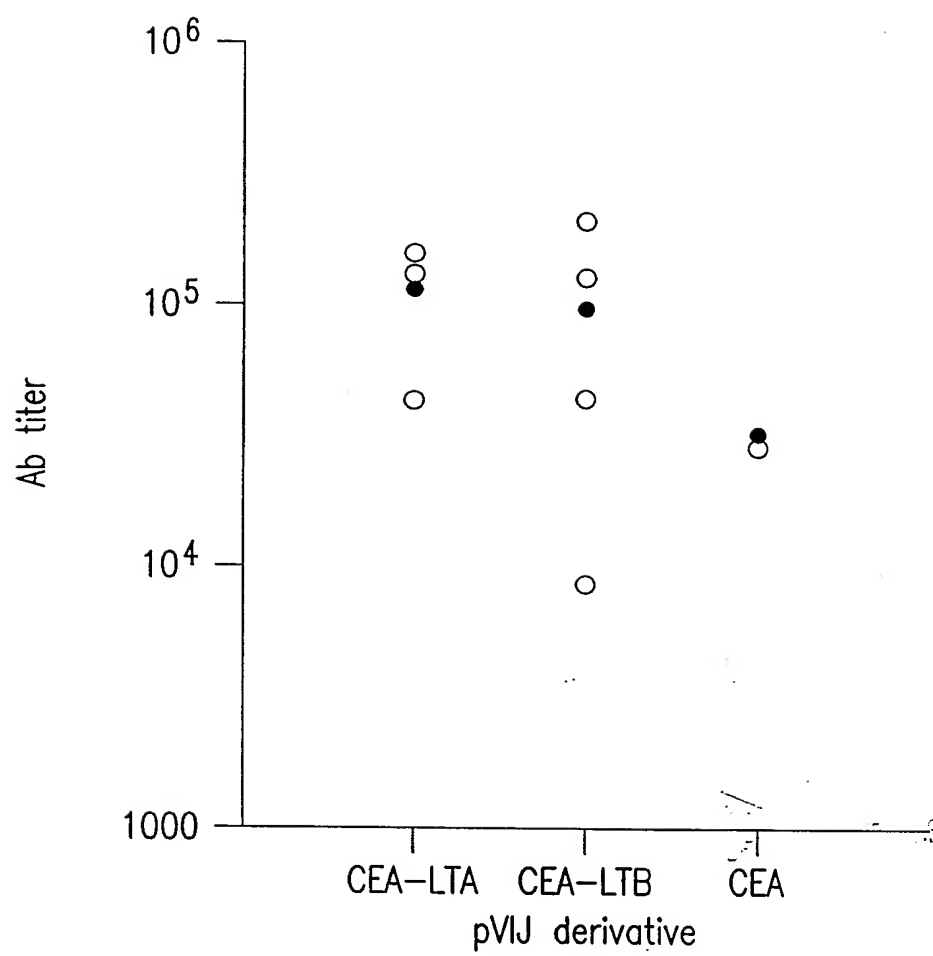


FIG. 11

28/62

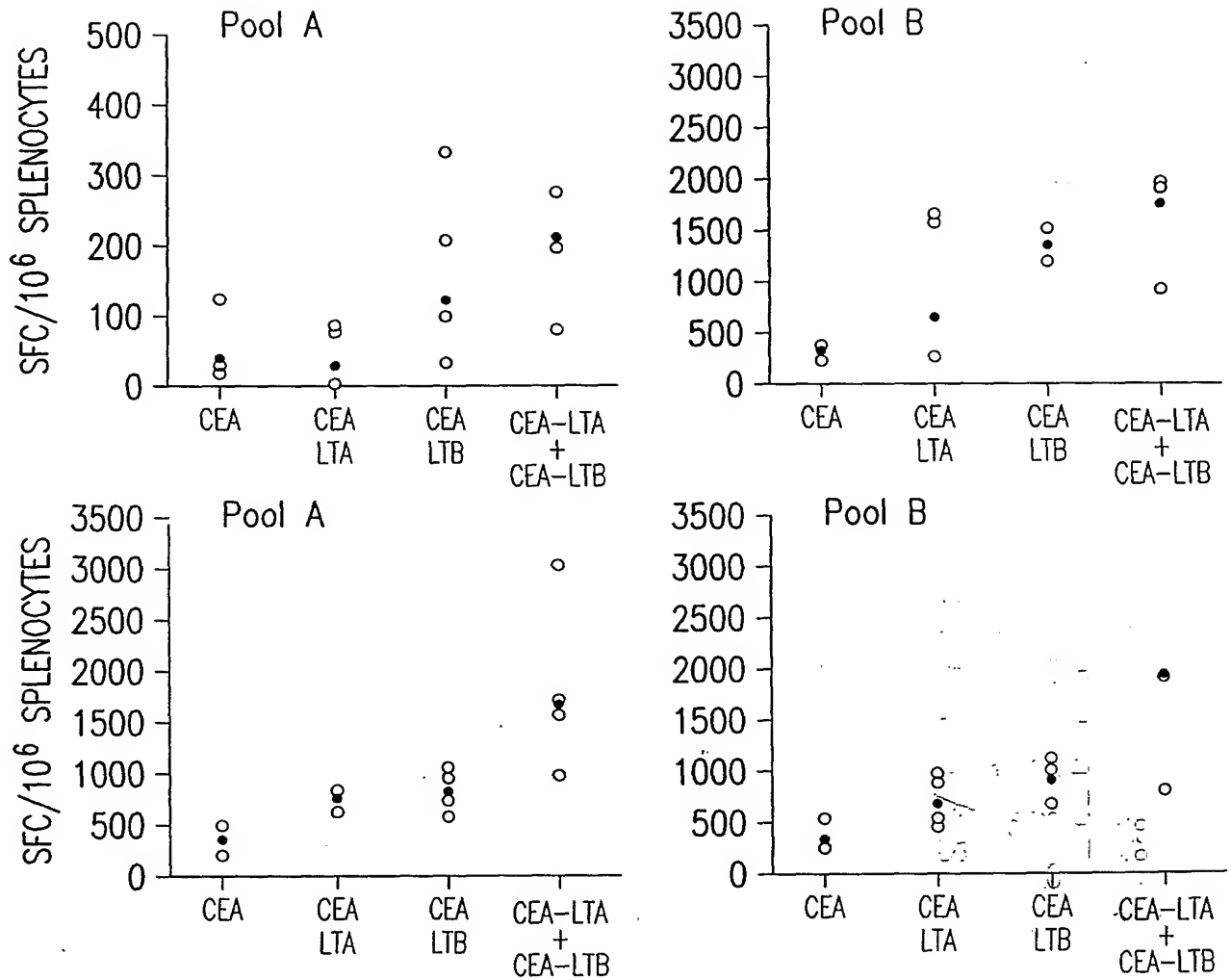


FIG. 1.2

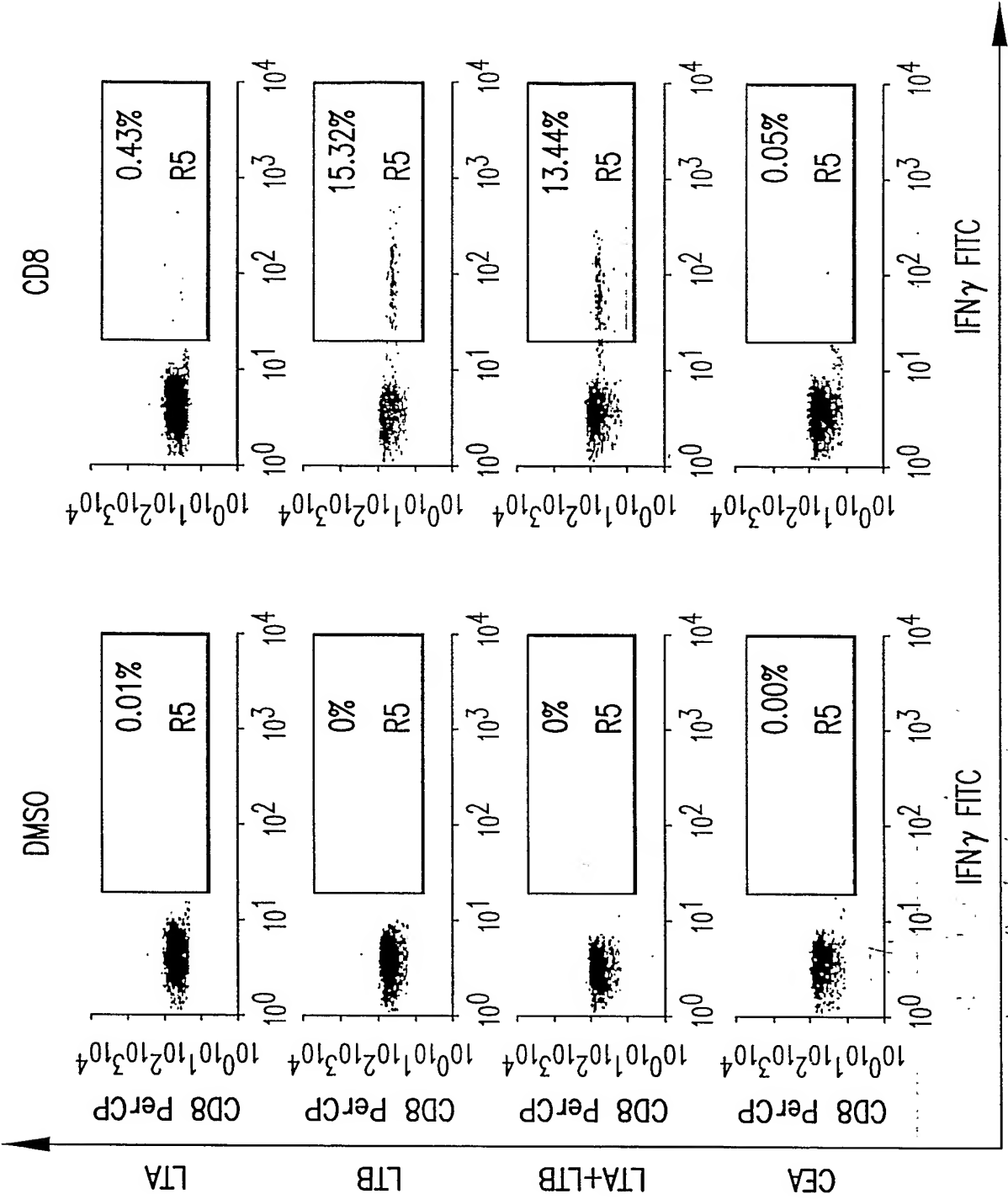


FIG.13

30/62

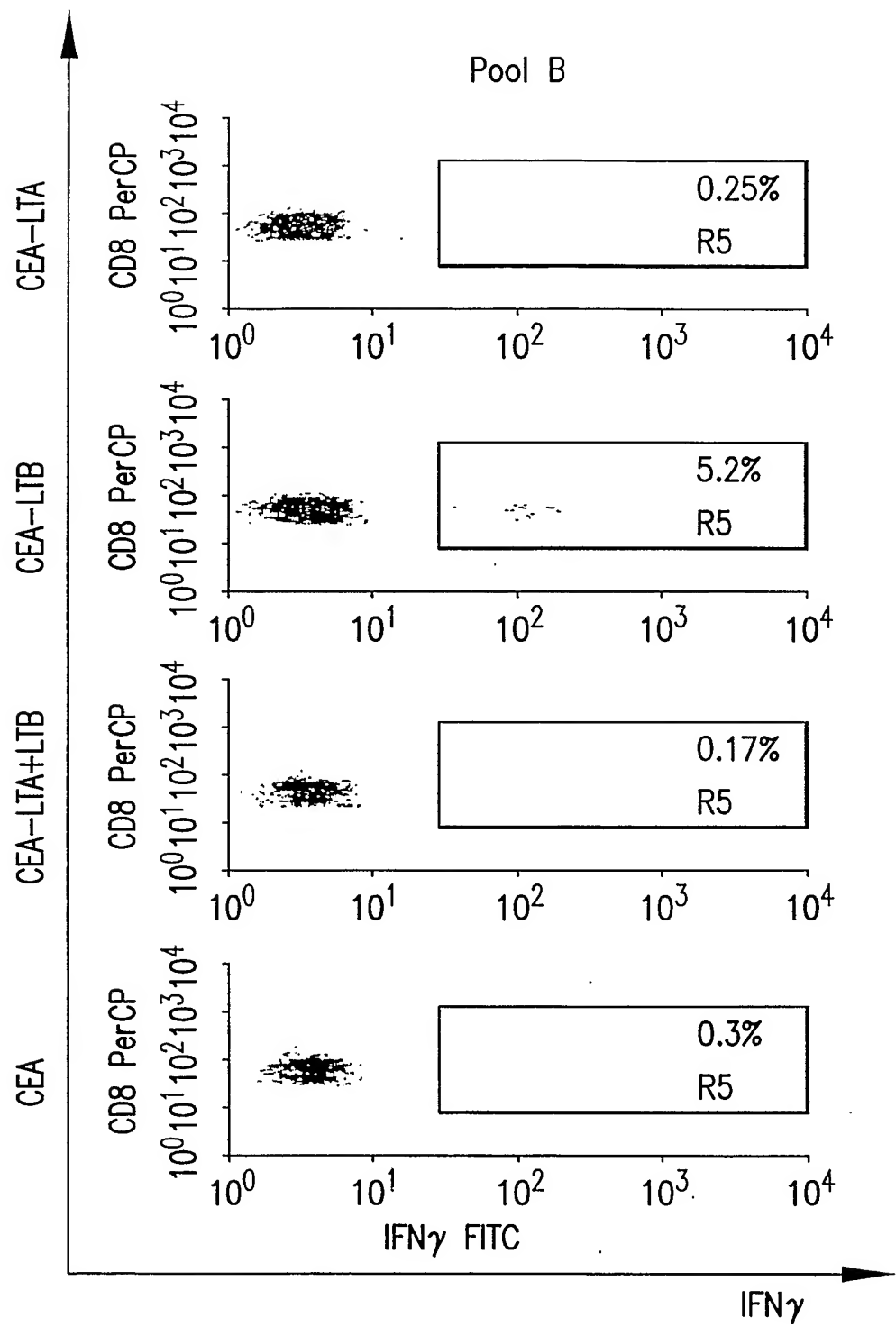


FIG. 14A

31/62

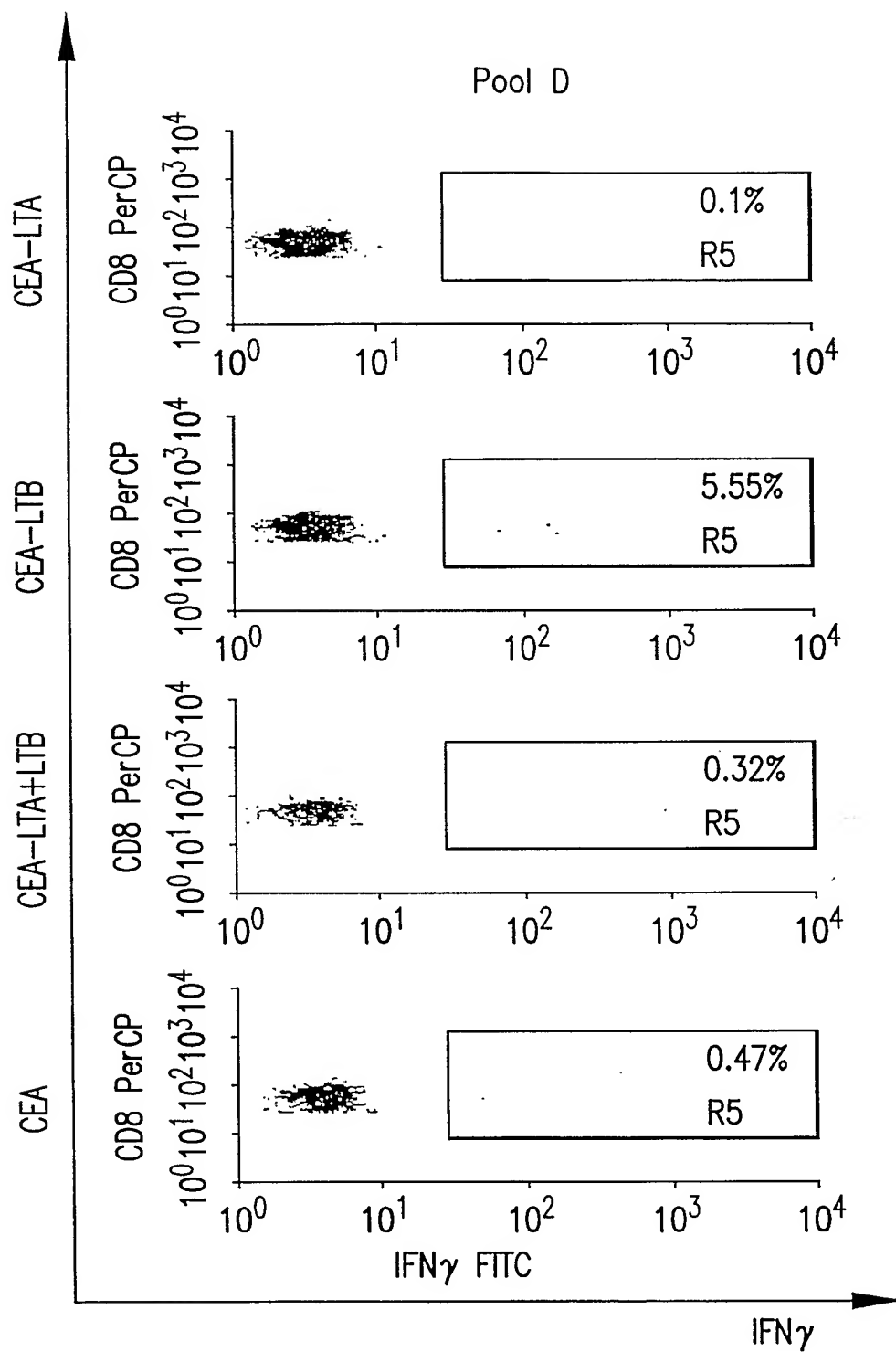


FIG.14B

32/62

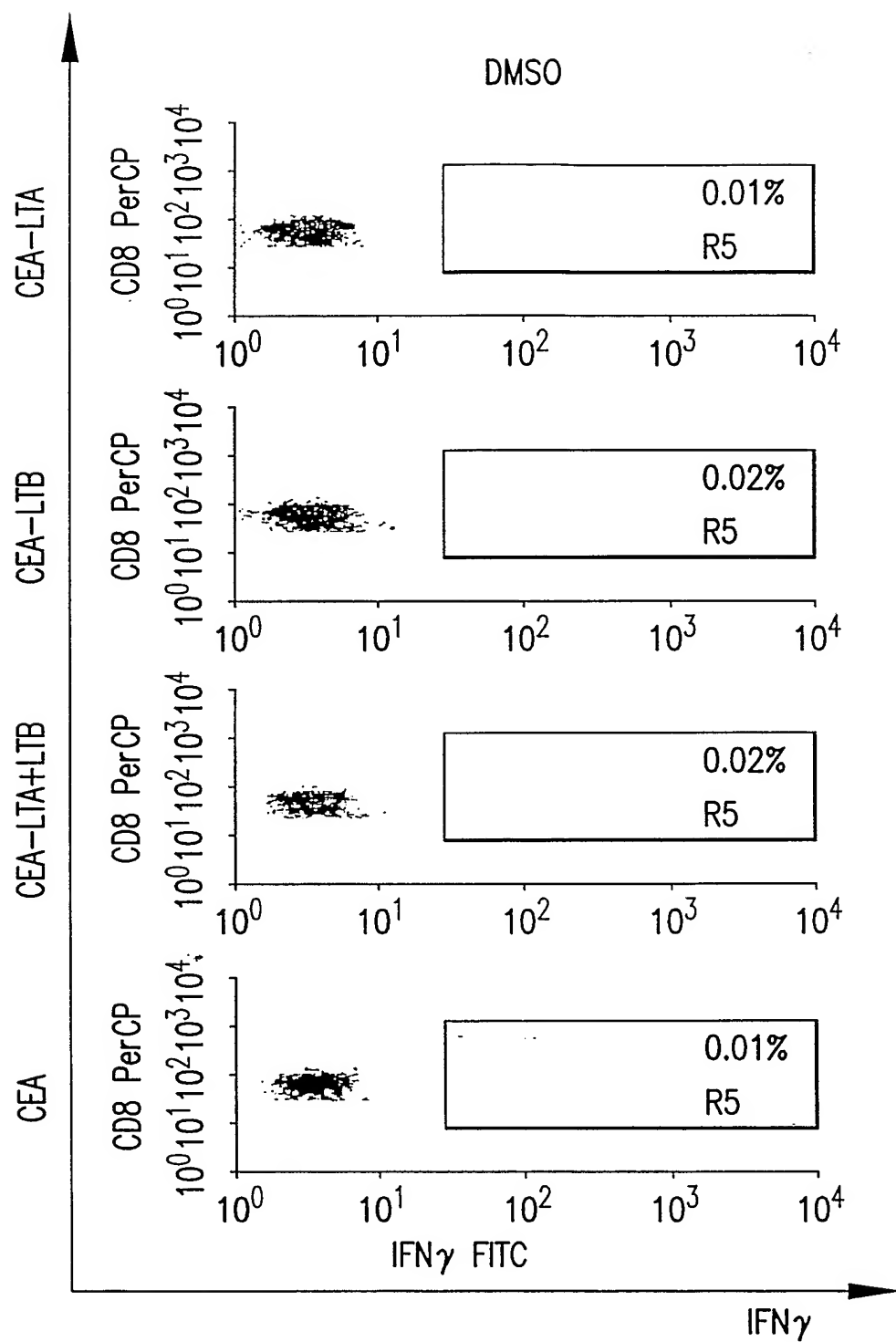


FIG. 14C



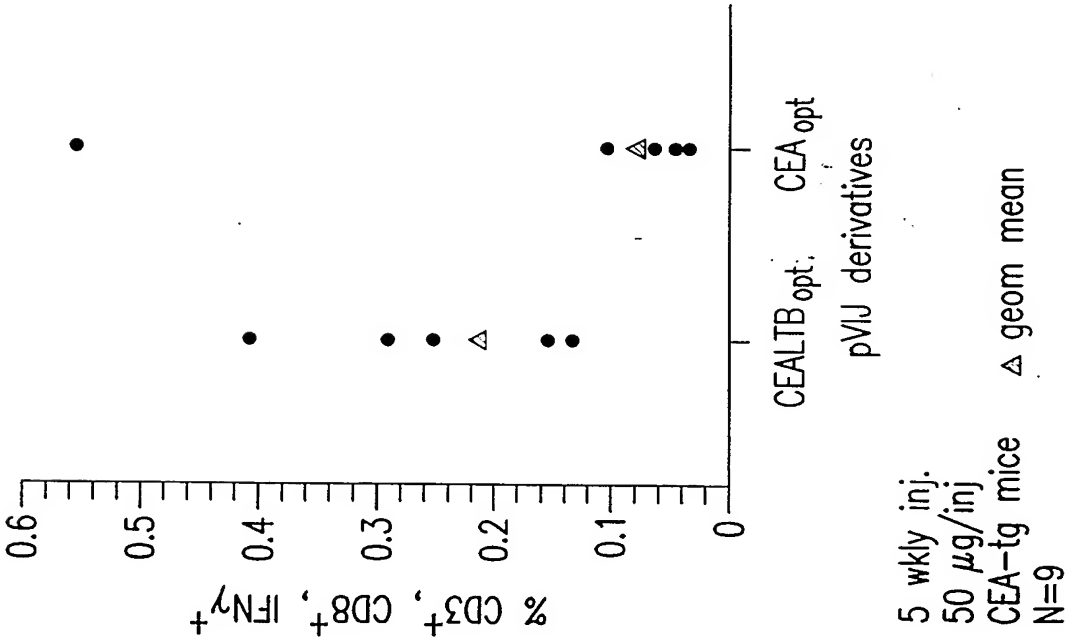


FIG.15A

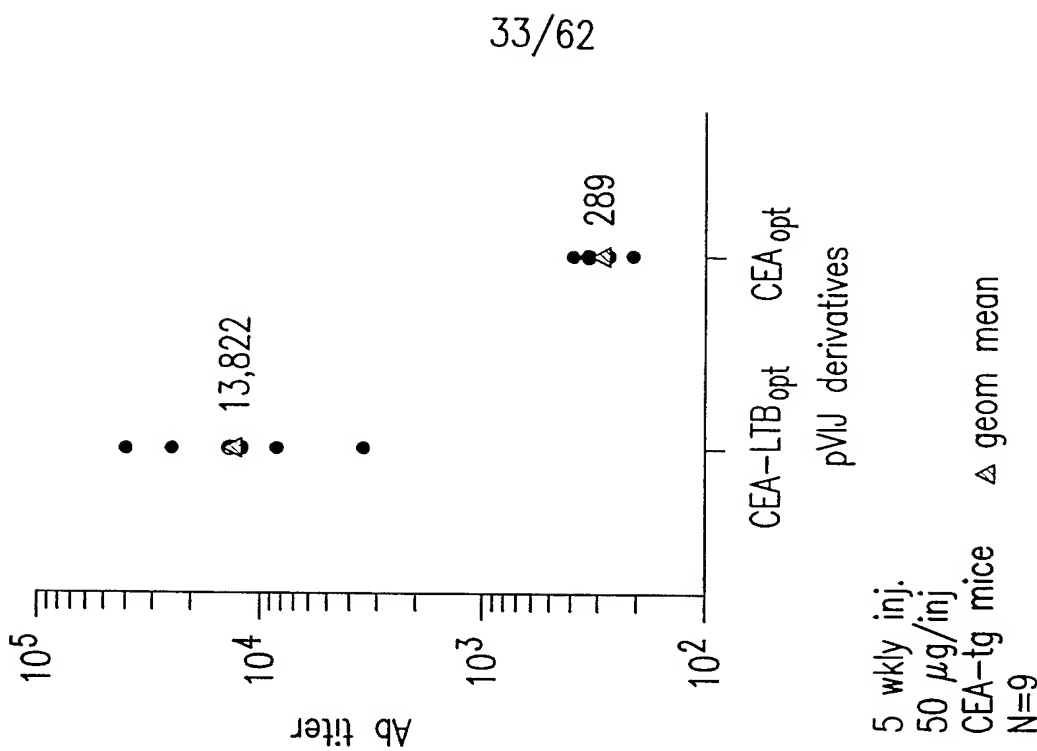


FIG.15B

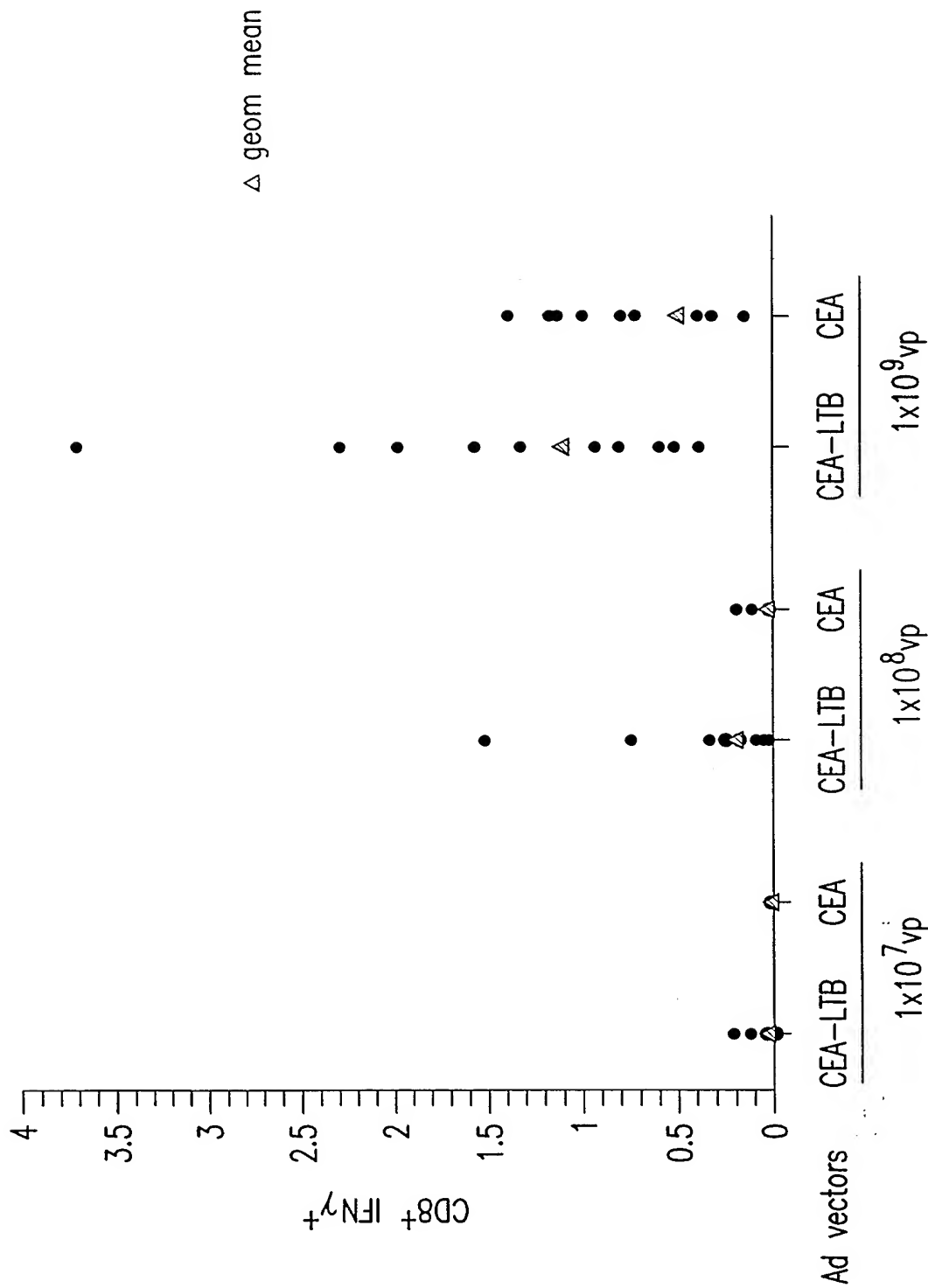


FIG.16

35/62

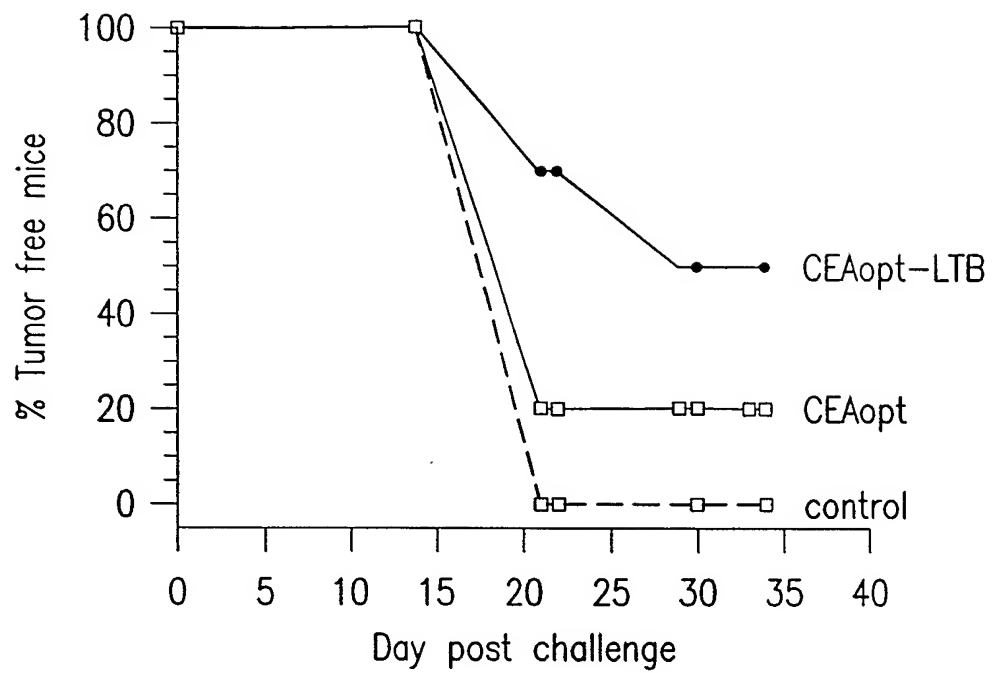


FIG.17A

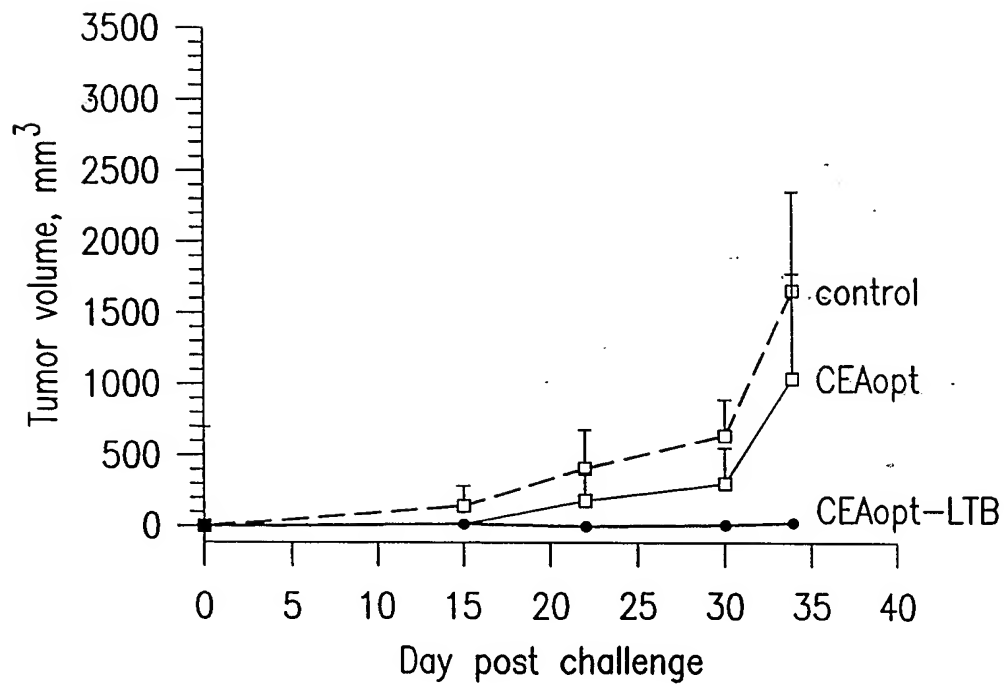


FIG.17B

36/62

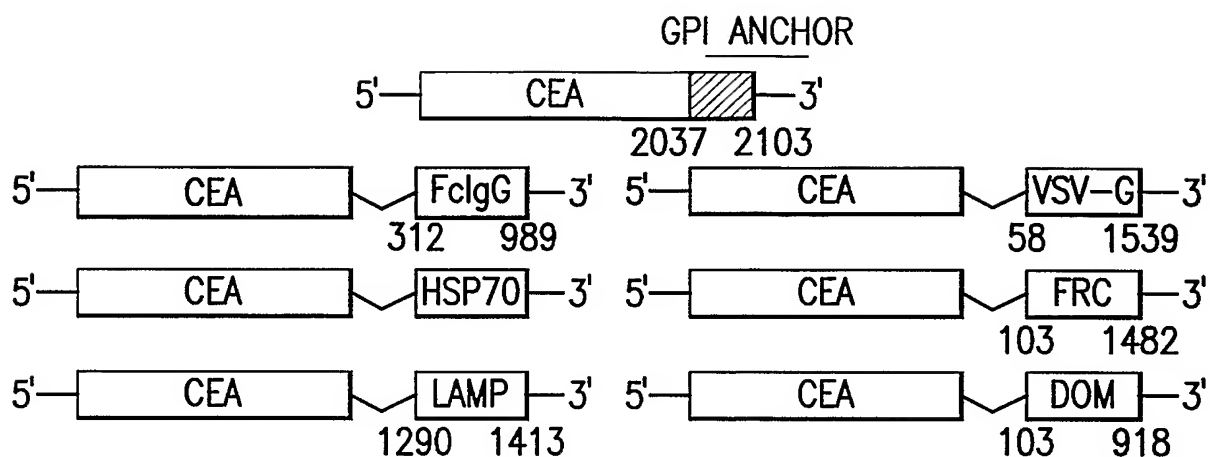


FIG.18A

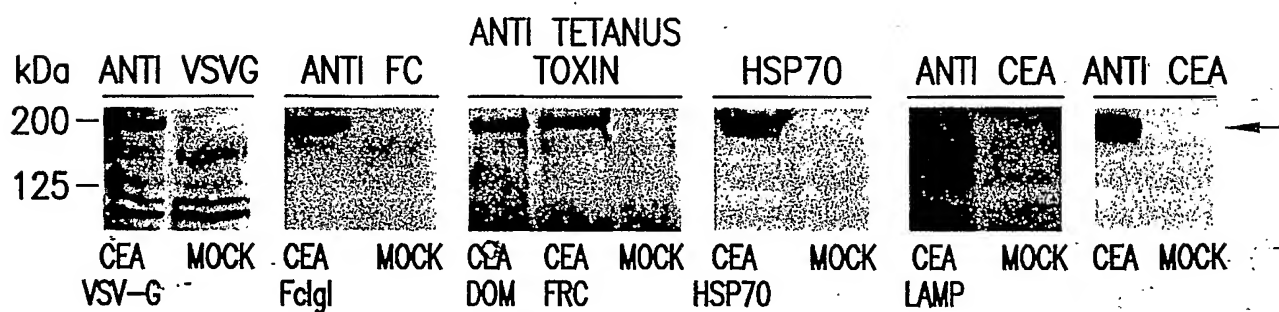


FIG.18B

37/62

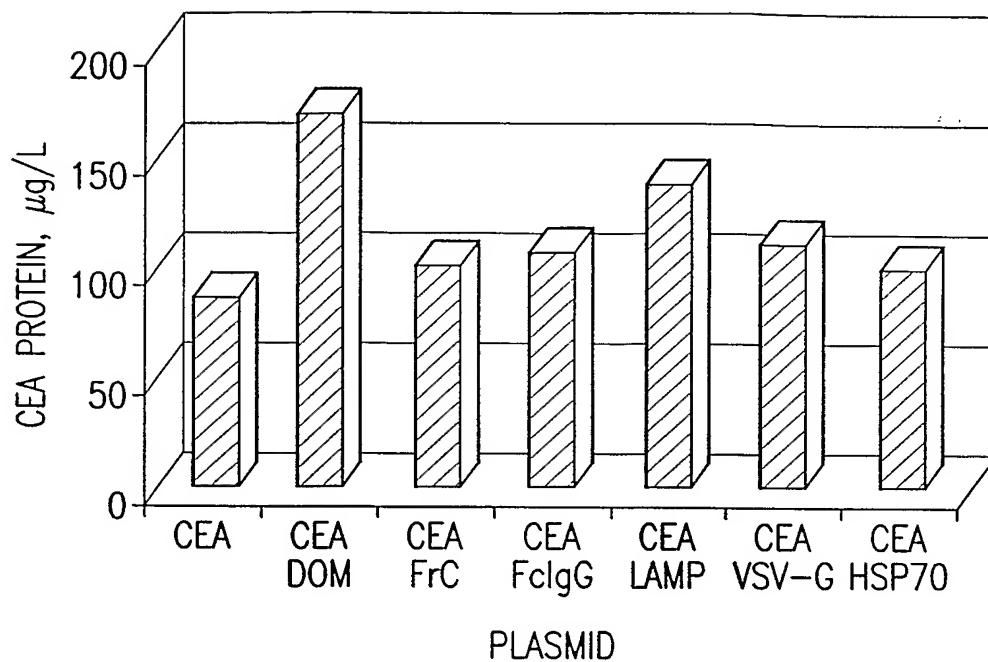


FIG.19A

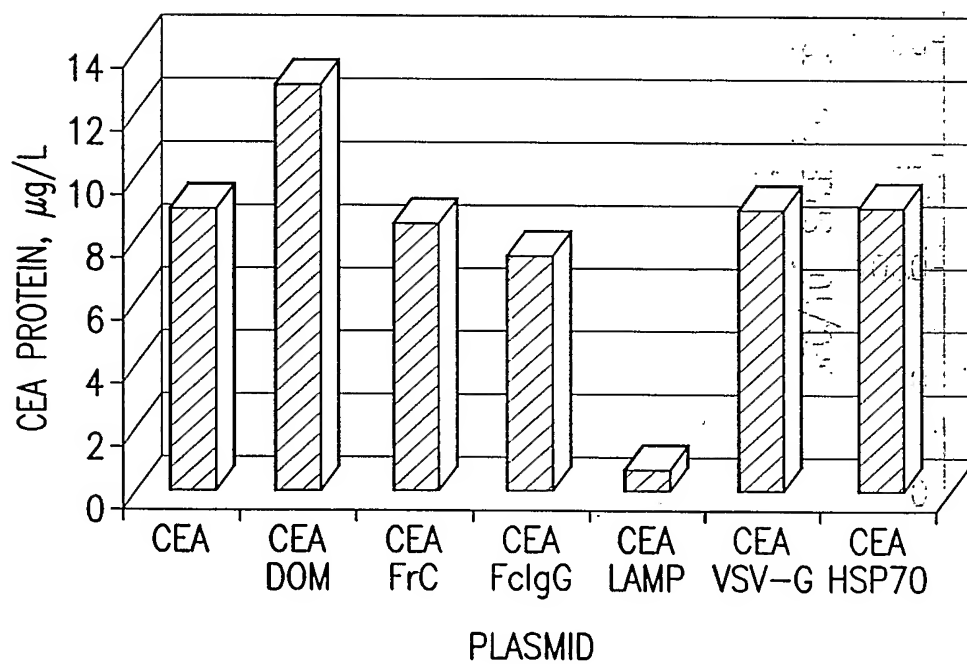


FIG.19B

38/62

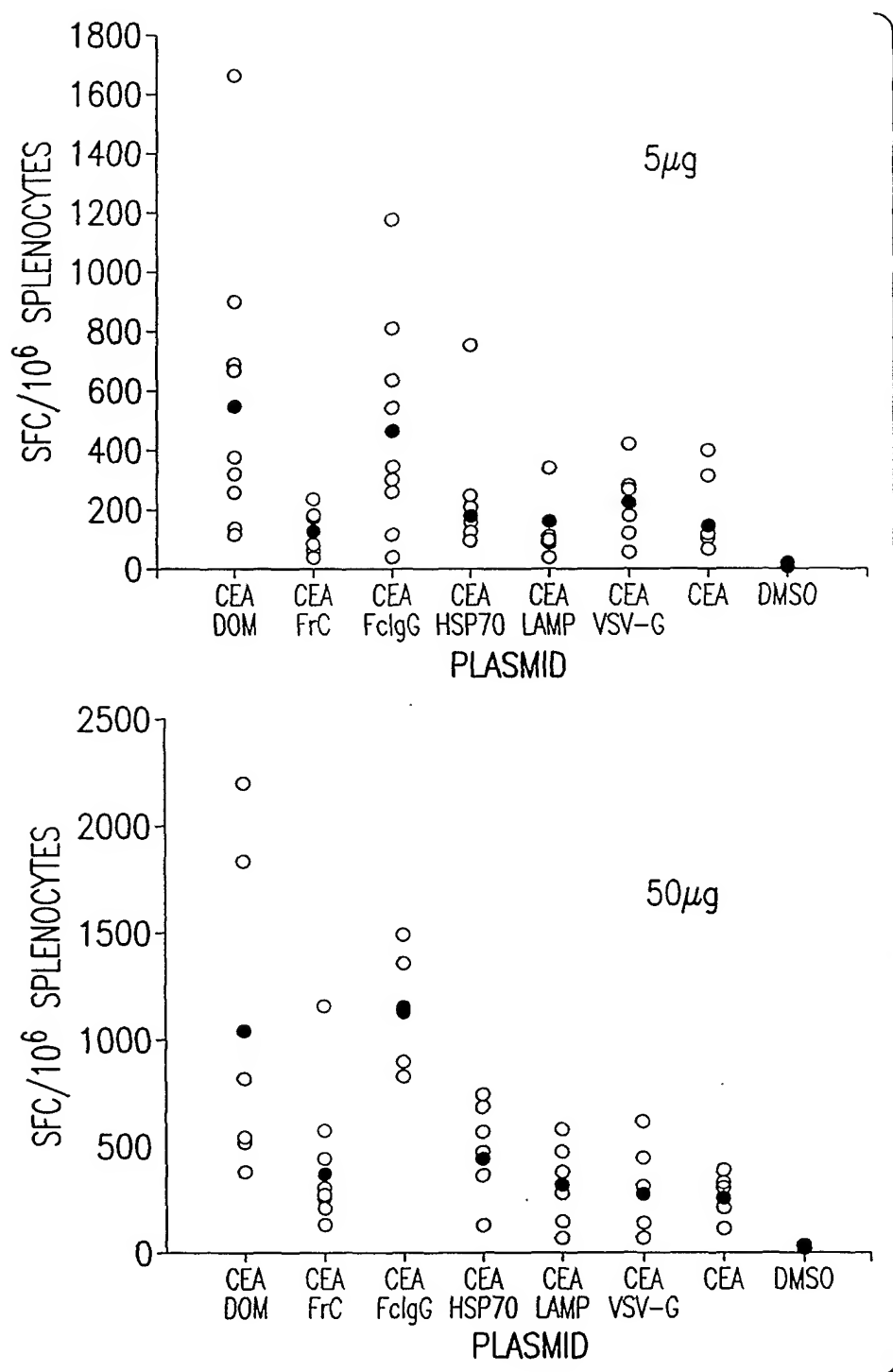


FIG. 20A

39/62

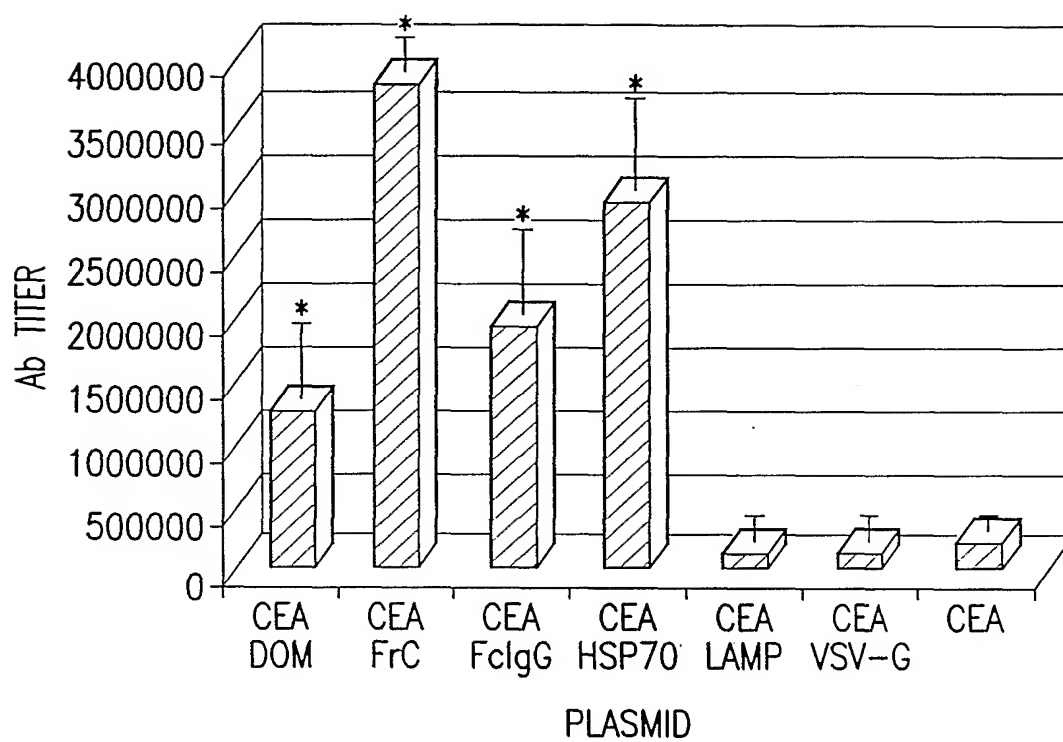


FIG.20B

40/62

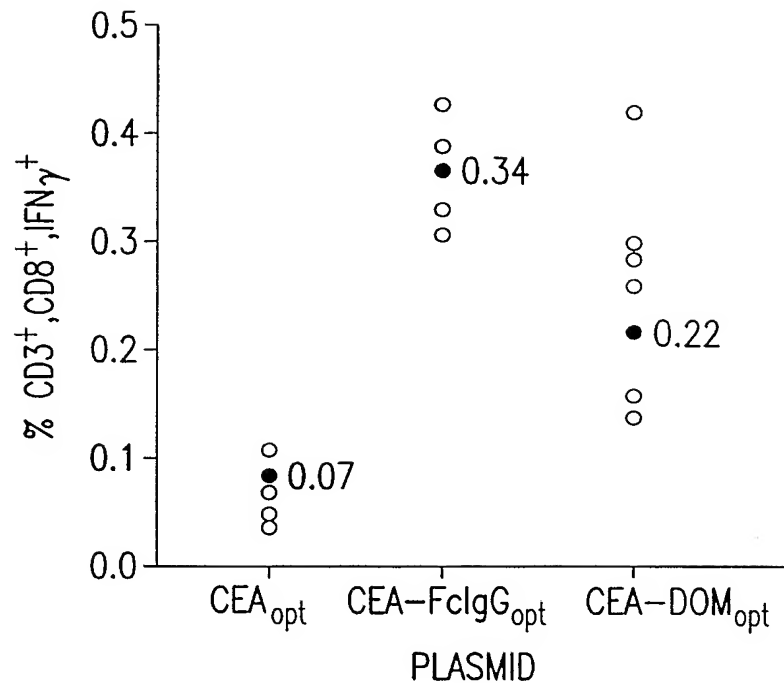


FIG.21A

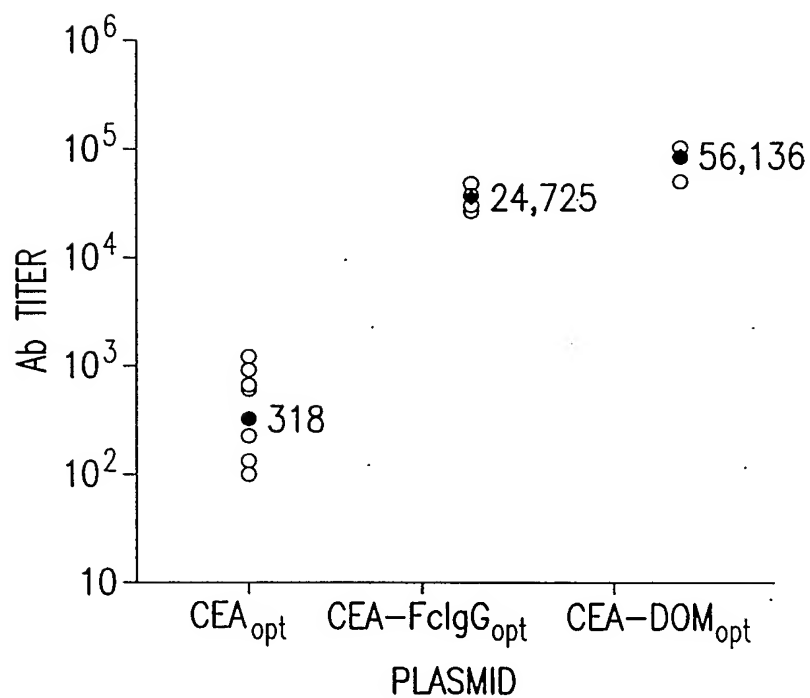


FIG.21B



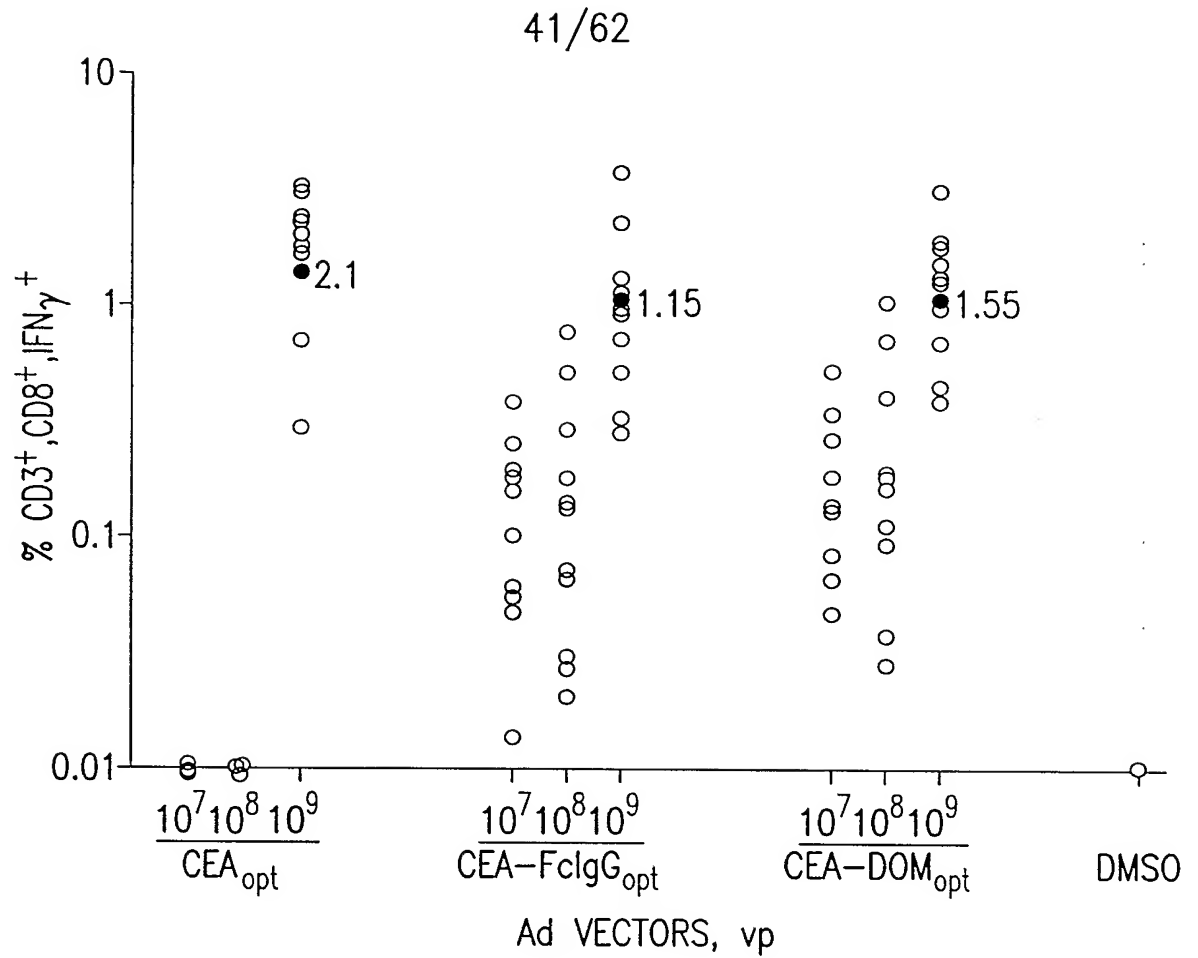


FIG.21C

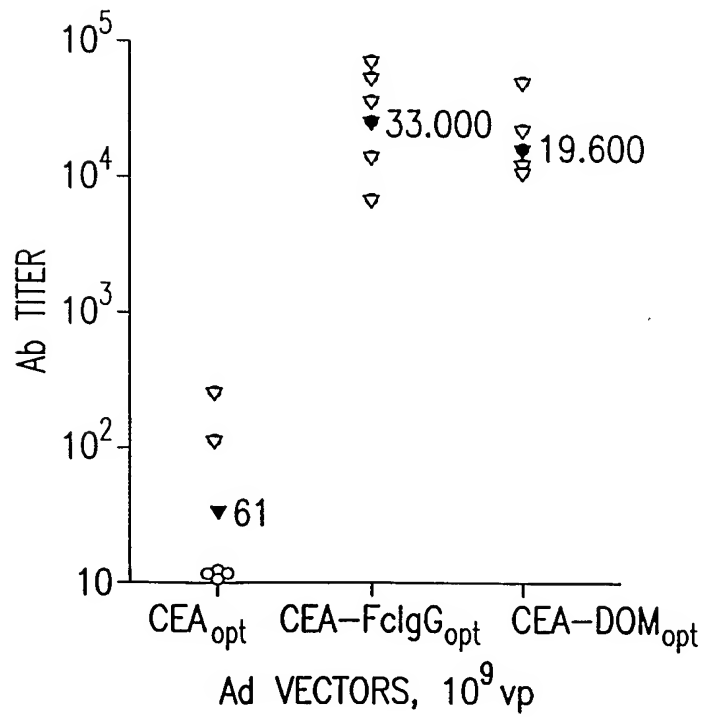


FIG.21D

42/62

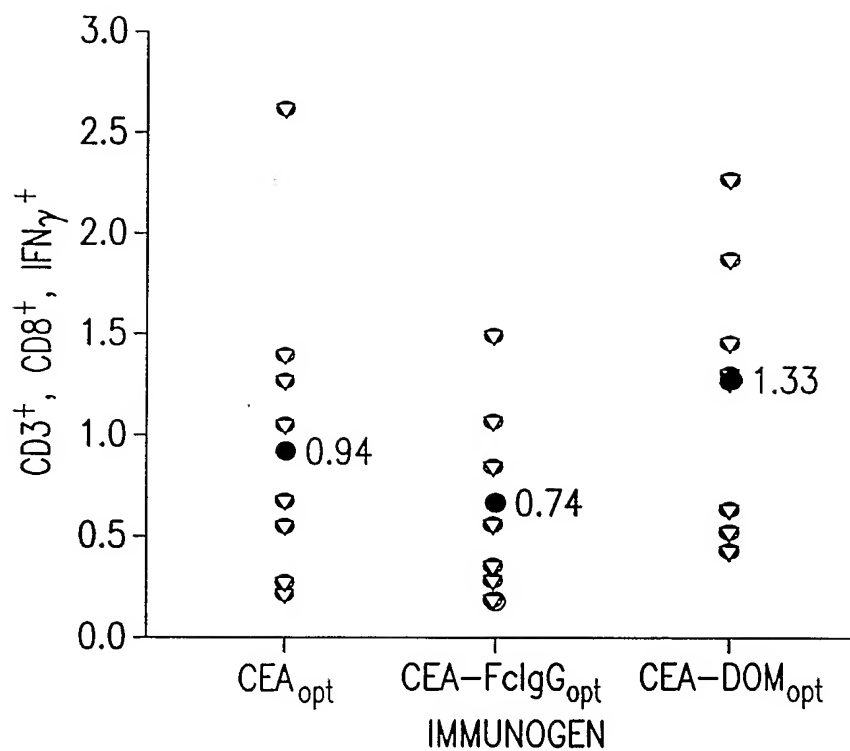


FIG.22A

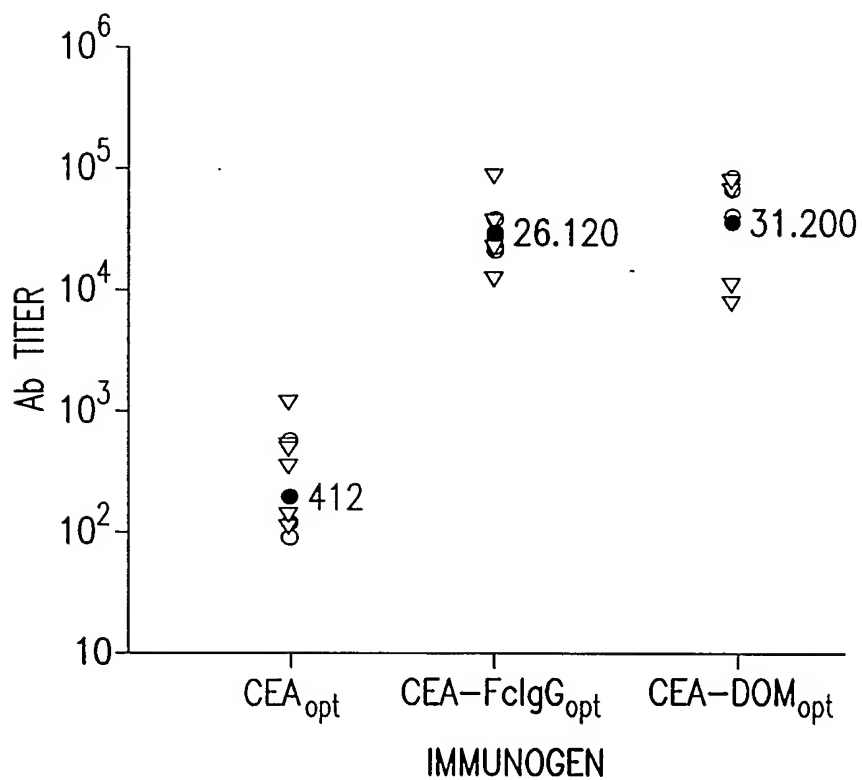


FIG.22B

43/62

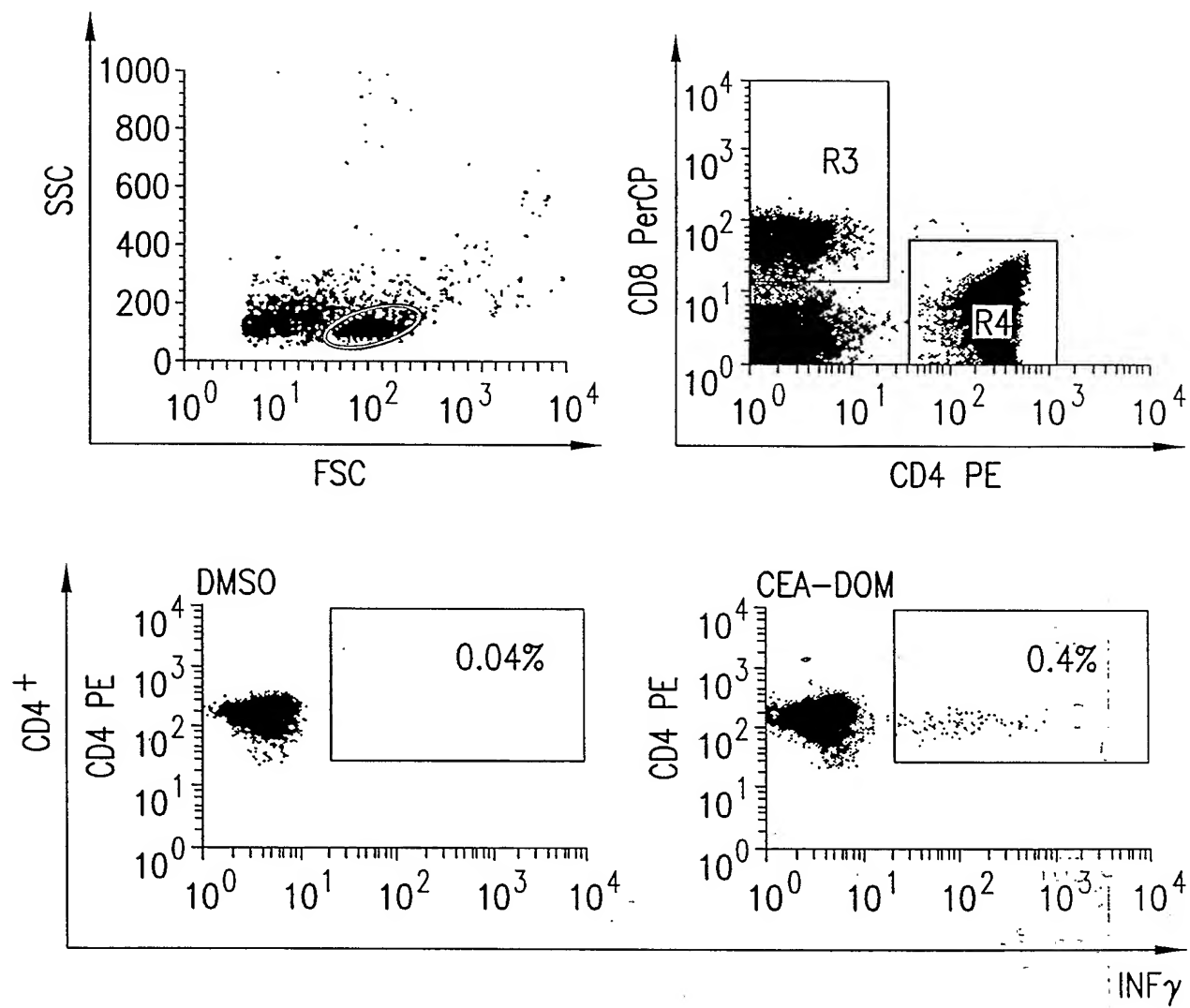


FIG.23

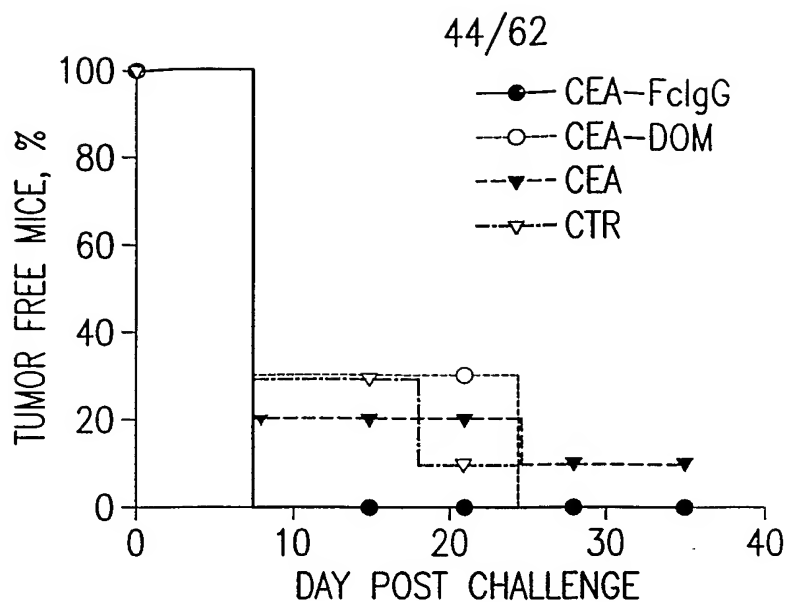


FIG. 24A

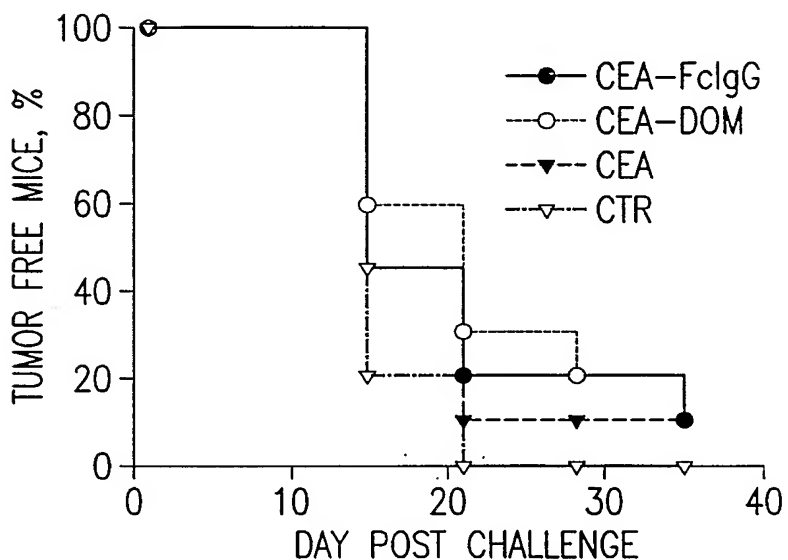


FIG. 24B

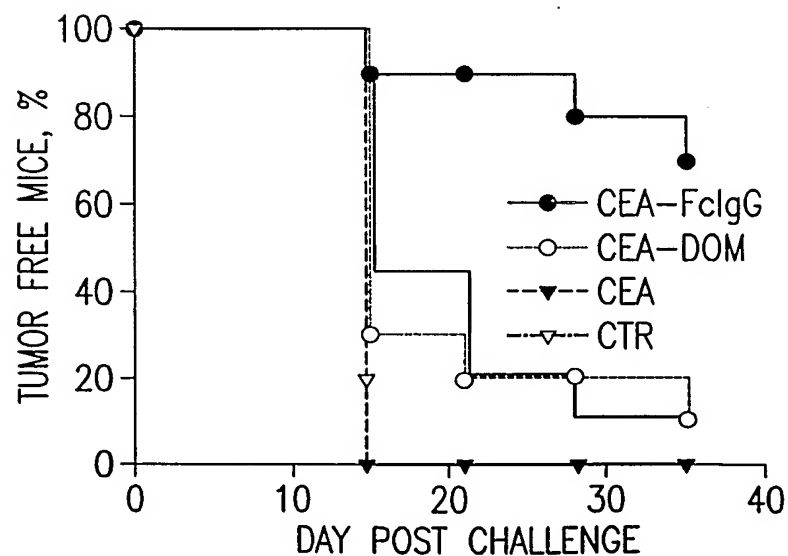


FIG. 24C

45/62

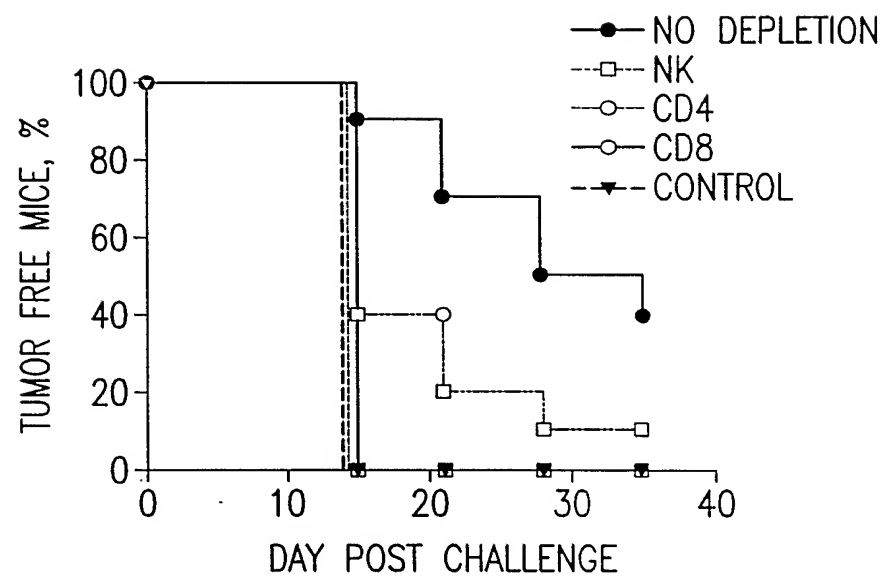


FIG.25

46/62

M E S P S A P P H R W C I P W Q R ·  
 1 ATGGAGAGCC CCAGCGCCCC CCCCCACCGC TGGTGCATCC CCTGGCAGCG  
 · L L L T A S L L T F W N P P T T A  
 51 CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG  
 · K L T I E S T P F N V A E G K E  
 101 CCAAGCTGAC CATCGAGAGC ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG  
 V L L L V H N L P Q H L F G Y S W ·  
 151 GTGCTGCTGC TGGTGCACAA CCTGCCCCAG CACCTGTTCC GCTACAGCTG  
 · Y K G E R V D G N R Q I I G Y V I  
 201 GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC GGCTACGTGA  
 · G T Q Q A T P G P A Y S G R E I  
 251 TCGGCACCCA GCAGGCCACC CCCGGCCCCG CCTACAGCGG CCGCGAGATC  
 I Y P N A S L L I Q N I I Q N D T ·  
 301 ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC  
 · G F Y T L H V I K S D L V N E E A  
 351 CGGCTTCTAC ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG  
 · T G Q F R V Y P E L P K P S I S  
 401 CCACCGGCCA GTTCCGCGTG TACCCCGAGC TGCCCAAGCC CAGCATCAGC  
 S N N S K P V E D K D A V A F T C ·  
 451 AGCAACAACA GCAAGCCCGT GGAGGACAAG GACGCCGTGG CCTTCACCTG  
 · E P E T Q D A T Y L W W V N N Q S  
 501 CGAGCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG AACAACCAGA  
 · L P V S P R L Q L S N G N R T L  
 551 GCCTGCCCCG GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCTG  
 T L F N V T R N D T A S Y K C E T ·  
 601 ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC  
 · Q N P V S A R R S D S V I L N V L  
 651 CCAGAACCCC GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC  
 · Y G P D A P T I S P L N T S Y R ·  
 701 TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CAGTACCGC  
 S G E N L N L S C H A A S N P P A ·  
 751 AGCGGCGAGA ACCTGAACCT GAGCTGCCAC GCCGCCAGCA ACCCCCCCGC  
 · Q Y S W F V N G T F Q Q S T Q E L  
 801 CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC  
 · F I P N I T V N N S G S Y T C Q  
 851 TGTTTCATCC CAACATCACC GTGAACAACA GCGGCAGCTA CACCTGCCAG  
 A H N S D T G L N R T T V T T I T ·  
 901 GCCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC  
 · V Y A E P P K P F I T S N N S N P  
 951 CGTGTACGCC GAGCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC  
 · V E D E D A V A L T C E P E I Q

FIG.26A-1

47/62

1001 CCGTGGAGGA CGAGGACGCC GTGGCCCTGA CCTGCGAGCC CGAGATCCAG  
       N T T Y L W W V N N Q S L P V S P .  
 1051 AACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGC CCGTGAGCCC  
       · R L Q L S N D N R T L T L L S V T .  
 1101 CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG CTGAGCGTGA  
       · R N D V G P Y E C G I Q N E L S  
 1151 CCCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC  
       V D H S D P V I L N V L Y G P D D .  
 1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCCGACGA  
       · P T I S P S Y T Y Y R P G V N L S .  
 1251 CCCCACCATC AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA  
       · L S C H A A S N P P A Q Y S W L  
 1301 GCCTGAGCTG CCACGCCGCC AGCAACCCCC CCGCCCAGTA CAGCTGGCTG  
       I D G N I Q Q H T Q E L F I S N I .  
 1351 ATCGACGGCA ACATCCAGCA GCACACCCAG GAGCTGTTCA TCAGCAACAT  
       · T E K N S G L Y T C Q A N N S A S .  
 1401 CACCGAGAAG AACAGCGGCC TGTACACCTG CCAGGCCAAC AACAGCGCCA  
       · G H S R T T V K T I T V S A E L  
 1451 GCGGCCACAG CCGCACCACC GTGAAGACCA TCACCGTGAG CGCCGAGCTG  
       P K P S I S S N N S K P V E D K D .  
 1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA  
       · A V A F T C E P E A Q N T T Y L W .  
 1551 CGCCGTGGCC TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT  
       · W V N G Q S L P V S P R L Q L S  
 1601 GGTGGGTGAA CGGCCAGAGC CTGCCCCTGA GCCCCGCCT GCAGCTGAGC  
       N G N R T L T L F N V T R N D A R .  
 1651 AACGGCAACC GCACCCTGAC CCTGTTCAAC GTGACCCGCA ACGACGCCCC  
       · A Y V C G I Q N S V S A N R S D P .  
 1701 CGCCTACGTG TCGGCATCC AGAACAGCGT GAGCGCCAAC CGCAGCGACC  
       · V T L D V L Y G P D T P I I S P .  
 1751 CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCAT CATCAGCGCC  
       P D S S Y L S G A N L N L S C H S .  
 1801 CCCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG  
       · A S N P S P Q Y S W R I N G I P Q .  
 1851 CGCCAGCAAC CCCAGCCCC AGTACAGCTG GCGCATCAAC GGCATCCCC  
       · Q H T Q V L F I A K I T P N N N  
 1901 AGCAGCACAC CCAGGTGCTG TTCATCGCCA AGATCACCCC CAACAACAAC  
       G T Y A C F V S N L A T G R N N S .  
 1951 GGCACCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG  
       · I V K S I T V S A S G T S R S T P .  
 2001 CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT AGAAGCACCC  
       · I P F S Y S K N L D C W V D N E

FIG.26A-2

48/62

2051 CCATCCCATT CAGCTACAGC AAGAACCTGG ACTGCTGGGT GGACAACGAG  
E D I D V I L K K S T I L N L D I ·  
2101 GAGGACATCG ACGTGATCCT GAAGAAGAGC ACCATCCTGA ACCTGGACAT  
· N N D I I S D I S G F N S S V I T ·  
2151 CAACAACGAC ATCATCAGCG ACATCAGCGG CTTCAACAGC AGCGTGATCA  
· Y P D A Q L V P G I N G K A I H  
2201 CCTACCCCGA CGCCCAGCTG GTGCCCGGCA TCAACGGCAA GGCCATCCAC  
L V N N E S S E V I V H K A M D I ·  
2251 CTGGTGAACA ACGAGAGCAG CGAGGTGATC GTGCACAAGG CCATGGACAT  
· E Y N D M F N N F T V S F W L R V ·  
2301 CGAGTACAAC GACATGTTCA ACAACTTCAC CGTGAGCTTC TGGCTGAGAG  
· P K V S A S H L E Q Y G T N E Y  
2351 TGCCTAAGGT GAGCGCCAGC CACCTGGAGC AGTACGGCAC CAACGAGTAC  
S I I S S M K K H S L S I G S G W ·  
2401 AGCATCATCA GCAGCATGAA GAAGCACAGC CTGAGCATCG GCAGCGGCTG  
· S V S L K G N N L I W T L K D S A ·  
2451 GAGCGTGAGC CTGAAGGGCA ACAACCTCAT CTGGACCCTG AAGGATAGCG  
· G E V R Q I T F R D L P D K F N  
2501 CCGGAGAGGT GAGACAGATC ACCTTCAGAG ACCTGCCCCGA CAAGTTCAAT  
A Y L A N K W V F I T I T N D R L ·  
2551 GCCTACCTGG CCAACAAGTG GGTGTTTCATC ACCATCACCA ACGACAGACT  
· S S A N L Y I N G V L M G S A E I ·  
2601 GAGCAGCGCC AACCTGTACA TCAACGGCGT GTCATGGGC AGCGCCGAGA  
· T G L G A I R E D N N I T L K L  
2651 TCACCGGCCT GGGCGCCATC AGAGAGGACA ACAACATCAC CCTGAAGCTG  
D R C N N N N Q Y V S I D K F R I ·  
2701 GACAGATGCA ACAACAACAA CCAGTACGTG AGCATCGACA AGTTCCGGAT  
· F C K A L N P K E I E K L Y T S Y ·  
2751 CTTCTGCAAG GCCCTGAACC CCAAGGAGAT CGAGAAGCTG TACACCAGCT  
· L S I T F L R D F W G N P L R Y  
2801 ACCTGAGCAT CACCTTCCTG AGAGACTTCT GGGGCAACCC CCTGAGATAC  
D T \* (SEQ ID NO:45)  
2851 GACACCTAG (SEQ ID NO:21)

FIG.26A-3



49/62

M E S P S A P P H R W C I P W Q R L L L  
 1 ATGGAGAGCC CCAGCGCCCC CCCCACCGC TGGTGCATCC CCTGGCAGCG CCTGCTGCTG  
 T A S L L T F W N P P T T A K L T I E S  
 61 ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG CCAAGCTGAC CATCGAGAGC  
 T P F N V A E G K E V L L L V H N L P Q  
 121 ACCCCCTTCA ACGTGGCCGA GGGCAAGGAG GTGCTGCTGC TGGTGCACAA CCTGCCCCAG  
 H L F G Y S W Y K G E R V D G N R Q I I  
 181 CACCTGTTCG GCTACAGCTG GTACAAGGGC GAGCGCGTGG ACGGCAACCG CCAGATCATC  
 G Y V I G T Q Q A T P G P A Y S G R E I  
 241 GGCTACGTGA TCGGCACCCA GCAGGCCACC CCCGGCCCCG CCTACAGCGG CCGCGAGATC  
 I Y P N A S L L I Q N I I Q N D T G F Y  
 301 ATCTACCCCA ACGCCAGCCT GCTGATCCAG AACATCATCC AGAACGACAC CGGCTTCTAC  
 T L H V I K S D L V N E E A T G Q F R V  
 361 ACCCTGCACG TGATCAAGAG CGACCTGGTG AACGAGGAGG CCACCGGCCA GTTCCGCGTG  
 Y P E L P K P S I S S N N S K P V E D K  
 421 TACCCCGAGC TGCCCAAGCC CAGCATCAGC AGCAACAACA GCAAGCCCGT GGAGGACAAG  
 D A V A F T C E P E T Q D A T Y L W W V  
 481 GACGCCGTGG CCTTCACCTG CGAGCCCCGAG ACCCAGGACG CCACCTACCT GTGGTGGGTG  
 N N Q S L P V S P R L Q L S N G N R T L  
 541 AACAAACCAGA GCCTGCCCCG GAGCCCCCGC CTGCAGCTGA GCAACGGCAA CCGCACCCCTG  
 T L F N V T R N D T A S Y K C E T Q N P  
 601 ACCCTGTTCA ACGTGACCCG CAACGACACC GCCAGCTACA AGTGCGAGAC CCAGAACCCC  
 V S A R R S D S V I L N V L Y G P D A P  
 661 GTGAGCGCCC GCCGCAGCGA CAGCGTGATC CTGAACGTGC TGTACGGCCC CGACGCCCCC  
 T I S P L N T S Y R S G E N L N L S C H  
 721 ACCATCAGCC CCCTGAACAC CAGCTACCGC AGCGGCGAGA ACCTGAACCT GAGCTGCCAC  
 A A S N P P A Q Y S W F V N G T F Q Q S  
 781 GCCGCCAGCA ACCCCCCCGC CCAGTACAGC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC  
 T Q E L F I P N I T V N N S G S Y T C Q  
 841 ACCCAGGAGC TGTTTCATCCC CAACATCACC GTGAACAACA GCGGCAGCTA CAGCTGCCAG  
 A H N S D T G L N R T T V T T I T V Y A  
 901 GCCCACAACA GCGACACCGG CCTGAACCGC ACCACCGTGA CCACCATCAC CGTGTACGCC  
 E P P K P F I T S N N S N P V E D E D A  
 961 GAGCCCCCA AGCCCTTCAT CACCAGCAAC AACAGCAACC CCGTGGAGGA CGAGGACGCC  
 V A L T C E P E I Q N T T Y L W W V N N  
 1021 GTGGCCCTGA CCTGCGAGCC CGAGATCCAG AACACCACCT ACCTGTGGTG GGTGAACAAC  
 Q S L P V S P R L Q L S N D N R T L T L  
 1081 CAGAGCCTGC CCGTGAGCCC CCGCCTGCAG CTGAGCAACG ACAACCGCAC CCTGACCCTG  
 L S V T R N D V G P Y E C G I Q N E L S  
 1141 CTGAGCGTGA CCCGCAACGA CGTGGGCCCC TACGAGTGCG GCATCCAGAA CGAGCTGAGC  
 V D H S D P V I L N V L Y G P D D P T I

FIG.27A-1

50/62

1201 GTGGACCACA GCGACCCCGT GATCCTGAAC GTGCTGTACG GCCCCGACGA CCCCACCATC  
       S P S Y T Y Y R P G V N L S L S C H A A  
 1261 AGCCCCAGCT ACACCTACTA CCGCCCCGGC GTGAACCTGA GCCTGAGCTG CCACGCCGCC  
       S N P P A Q Y S W L I D G N I Q Q H T Q  
 1321 AGCAACCCCC CCGCCAGTA CAGCTGGCTG ATCGACGGCA ACATCCAGCA GCACACCCAG  
       E L F I S N I T E K N S G L Y T C Q A N  
 1381 GAGCTGTTCA TCAGCAACAT CACCGAGAAG AACAGCGGCC TGTACACCTG CCAGGCCAAC  
       N S A S G H S R T T V K T I T V S A E L  
 1441 AACAGCGCCA GCGGCCACAG CCGCACCACC GTGAAGACCA TCACCGTGAG CGCCGAGCTG  
       P K P S I S S N N S K P V E D K D A V A  
 1501 CCCAAGCCCA GCATCAGCAG CAACAACAGC AAGCCCGTGG AGGACAAGGA CGCCGTGGCC  
       F T C E P E A Q N T T Y L W W V N G Q S  
 1561 TTCACCTGCG AGCCCGAGGC CCAGAACACC ACCTACCTGT GGTGGGTGAA CGGCCAGAGC  
       L P V S P R L Q L S N G N R T L T L F N  
 1621 CTGCCCCGTA GCCCCCGCCT GCAGCTGAGC AACGGCAACC GCACCCTGAC CCTGTTCAAC  
       V T R N D A R A Y V C G I Q N S V S A N  
 1681 GTGACCCGCA ACGACGCCCG CGCCTACGTG TCGGGCATCC AGAACAGCGT GAGCGCCAAC  
       R S D P V T L D V L Y G P D T P I I S P  
 1741 CGCAGCGACC CCGTGACCCT GGACGTGCTG TACGGCCCCG ACACCCCAT CATCAGCCCC  
       P D S S Y L S G A N L N L S C H S A S N  
 1801 CCCGACAGCA GCTACCTGAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG CGCCAGCAAC  
       P S P Q Y S W R I N G I P Q Q H T Q V L  
 1861 CCCAGCCCCC AGTACAGCTG GCGCATCAAC GGCATCCCCC AGCAGCACAC CCAGGTGCTG  
       F I A K I T P N N N G T Y A C F V S N L  
 1921 TTCATCGCCA AGATCACCCC CAACAACAAC GGCACCTACG CCTGCTTCGT GAGCAACCTG  
       A T G R N N S I V K S I T V S A S G T S  
 1981 GCCACCGGCC GCAACAACAG CATCGTGAAG AGCATCACCG TGAGCGCCAG CGGCACCTCT  
       R K T H T C P P C P A P E L L G G P S V  
 2041 AGAAAGACCC ACACCTGCCC CCCTTGCCCT GCCCCTGAGC TGCTGGGCGG ACCCAGCGTG  
       F L F P P K P K D T L M I S R T P E V T  
 2101 TTCCTGTTCC CCCCCAAGCC TAAGGACACC CTCATGATCA GCAGAACCCC CGAGGTGACC  
       C V V V D V S H E D P E V K F N W Y V D  
 2161 TCGTGGTGG TGGACGTGAG CCACGAGGAT CCCGAGGTGA AGTTCAACTG GTACGTGGAC  
       G V E V H N A K T K P R E E Q Y N S T Y  
 2221 GCGTGGAGG TGCACAATGC CAAGACCAAG CCCAGAGAGG AGCAGTACAA CAGCACCTAC  
       R V V S V L T V L H Q D W L N G K E Y K  
 2281 AGAGTGGTGA GCGTGCTCAC CGTGCTGCAC CAGGATTGGC TGAACGGCAA GGAGTACAAG  
       C K V S N K A L P A P I E K T I S K A K  
 2341 TGCAAGGTGA GCAACAAGGC CCTGCCTGCC CCCATCGAGA AAACCATCAG CAAGGCCAAG  
       G Q P R E P Q V Y T L P P S R D E L T K

FIG.27A-2

51/62

2401 GGCCAGCCCA GAGAGCCCA GGTGTACACC CTGCCCCCTA GCAGAGATGA GTTGACCAAG  
N Q V S L T C L V K G F Y P S D I A V E  
2461 AACCAGGTGA GCCTCACATG CCTGGTGAAG GGCTTCTACC CCAGCGACAT CGCCGTGGAG  
W E S N G Q P E N N Y K T T P P V L D S  
2521 TGGGAGAGCA ACGGCCAGCC CGAGAACAAC TACAAGACCA CCCCCCTGT GCTGGACAGC  
D G S F F L Y S K L T V D K S R W Q Q G  
2581 GATGGCAGCT TCTTCCTGTA CAGCAAGCTC ACCGTGGACA AGAGCAGATG GCAGCAGGGC  
N V F S C S V M H E A L H N H Y T Q K S  
2641 AACGTGTTCA GCTGCAGCGT GATGCACGAG GCCCTGCACA ATCACTACAC CCAGAAGAGC  
L S L S P G K \* (SEQ ID NO:46)  
2701 CTGAGCCTGA GCCCCGGCAA GTAA (SEQ ID NO:25)

FIG.27A-3

52/62

```

atggagtctc cctcggcccc tccccacaga tggatcatcc cctggcagag gctcctgctc 60
acagcctcac ttctaacctt ctggaacccg cccaccactg ccaagctcac tattgaatcc 120
acgccgttca atgtcgcaga ggggaaggag gtgcttctac ttgtccacaa tctgccccag 180
catctttttg gctacagctg gtacaaagggt gaaagagtgg atggcaaccg tcaaattata 240
ggatatgtaa taggaactca acaagctacc ccagggcccg catacagtgg tcgagagata 300
atatacccca atgcatccct gctgatccag aacatcatcc agaatgacac aggattctac 360
accctacacg tcataaagtc agatcttgtg aatgaagaag caactggcca gttccgggta 420
taccgggagc tgcccaagcc ctccatctcc agcaacaact ccaaaccggt ggaggacaag 480
gatgctgtgg ccttcacctg tgaacctgag actcaggacg caacctacct gtggtgggta 540
aacaatcaga gcctcccggg cagtcccagg ctgcagctgt ccaatggcaa caggaccctc 600
actctattca atgtcacaag aaatgacaca gcaagctaca aatgtgaaac ccagaaccca 660
gtgagtgccg ggcgcagtga ttcagtcata ctgaatgtcc tctatggccc ggatgcccc 720
accatttccc ctctaaacac atcttacaga tcaggggaaa atctgaacct ctctgccac 780
gcagcctcta acccactgc acagtactct tggtttgtca atgggacttt ccagcaatcc 840
acccaagagc tctttatccc caacatcact gtgaataata gtggatccta tacgtgccaa 900
gcccataact cagacactgg cctcaatagg accacagtca cgacgatcac agtctatgca 960
gagccacca aacccttcat caccagcaac aactccaacc ccgtggagga tgaggatgct 1020
gtagccttaa cctgtgaacc tgagattcag aacacaacct acctgtggtg ggtaaataat 1080
cagagcctcc cgtcagtc caggctgcag ctgtccaatg acaacaggac cctcactcta 1140
ctcagtgta caaggaatga ttaggagacc tatgagtgtg gaatccagaa cgaattaagt 1200
gttgaccaca gcgaccagt catcctgaat gtcctctatg gccagacga cccaccatt 1260
tccccctcat acacctatta ccgtccagggt gtgaacctca gcctctcctg ccatgcagcc 1320
tctaaccac ctgcacagta ttcttggtg attgatggga acatccagca acacacaaa 1380
gagctcttta tctccaacat cactgagaag aacagcggac tctataacct ccaggccaat 1440
aactcagcca gtggccacag caggactaca gtcaagacaa tcacagtctc tgcggagctg 1500
cccaagccct ccatctccag caacaactcc aaaccctggg aggacaagga tgctgtggcc 1560
ttcacctgtg aacctgaggc tcagaacaca acctacctgt ggtgggtaaa tggtcagagc 1620
ctcccagtca gtcccagggt gcagctgtcc aatggcaaca ggaccctcac tctattcaat 1680
gtcacaagaa atgacgcaag agcctatgta tgtggaatcc agaactcagt gagtgcacaa 1740
cgcagtgacc cagtcaccct ggatgtcctc tatgggccgg acacccccat catttcccc 1800
ccagactcgt cttacctttc gggagcgaac ctcaacctct cctgccactc ggcctctaac 1860
ccatccccgc agtattcttg gcgtatcaat gggataccgc agcaacacac acaagttctc 1920
tttatcgcca aatcacgcc aaataataac gggacctatg cctgttttgt ctctaacttg 1980
gctactggcc gcaataattc catagtcaag agcatcacag tctctgcata tgggaact(SEQ. ID NO. 22)

```

FIG.28A

53/62

1 MESPSAPP<sup>5</sup>HR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVAEGKE  
51 VLLL VHNLPQ HLF GYSWYKG ERVDGNRQII GYVIGTQQAT PGPAYSGREI  
101 IYPNASLLIQ NIIQNDTGFY TLHVIKSDLV NEEATGQFRV YPELPKPSIS  
151 SNNSKPVEDK DAVAFTCEPE TQDATYLWWV NNQSLPVSPR LQLSNGNRTL  
201 TLFNVTRNDT ASYKCETQNP VSARRSDSVI LNVLYGPDAP TISPLNTSYR  
251 SGENLNLSCH AASNPPAQYS WFNVTGFQQS TQELFIPNIT VNNSGSYTCQ  
301 AHNSDTGLNR TTVTTITVYA EPPKPFITSN NSNPVEDEDA VALTCEPEIQ  
351 NTTYLWWVNN QSLPVSPRLQ LSNDNRTLTL LSVTRNDVGP YECGIQNELS  
401 VDHS DPVILN VLYGPDDPTI SPSYTYRPG VNLSLSCHAA SNPPAQYSWL  
451 IDGNIQQHTQ ELFISNITEK NSGLYTCQAN NSASGHSRTT VKTITVSAEL  
501 PKPSISSNNS KPVEDKDAVA FTCEPEAQNT TYLWWVNGQS LPVSPRLQLS  
551 NGNRTLTLFN VTRNDARAYV CGIQNSVSAN RSDPVTLDVL YGPDTPIIISP  
601 PDSSYLSGAN LNLSCHSASN PSPQYSWRIN GIPQQHTQVL FIAKITPNNN  
651 GTYACFVSNL ATGRNNSIVK SITVSASGT (SEQ ID NO:23)

FIG.28B

54/62

D S T P I P F S Y S K N L D C W V D N E  
1 GATTCAACAC CAATTCCATT TTCTTATTCT AAAAATCTGG ATTGTTGGGT TGATAATGAA  
E D I D V I L K K S T I L N L D I N N D  
61 GAAGATATAG ATGTTATATT AAAAAAGAGT ACAATTTTAA ATTTAGATAT TAATAATGAT  
I I S D I S G F N S S V I T Y P D A Q L  
121 ATTATATCAG ATATATCTGG GTTTAATTCA TCTGTAATAA CATATCCAGA TGCTCAATTG  
V P G I N G K A I H L V N N E S S E V I  
181 GTGCCCGGAA TAAATGGCAA AGCAATACAT TTAGTAAACA ATGAATCTTC TGAAGTTATA  
V H K A M D I E Y N D M F N N F T V S F  
241 GTGCATAAAG CTATGGATAT TGAATATAAT GATATGTTTA ATAATTTTAC CGTTAGCTTT  
W L R V P K V S A S H L E Q Y G T N E Y  
301 TGGTTGAGGG TTCCTAAAGT ATCTGCTAGT CATTTAGAAC AATATGGCAC AAATGAGTAT  
S I I S S M K K H S L S I G S G W S V S  
361 TCAATAATTA GCTCTATGAA AAAACATAGT CTATCAATAG GATCTGGTTG GAGTGTATCA  
L K G N N L I W T L K D S A G E V R Q I  
421 CTTAAAGGTA ATAACCTAAT ATGGACTTTA AAAGATTCCG CGGGAGAAGT TAGACAAATA  
T F R D L P D K F N A Y L A N K W V F I  
481 ACTTTTAGGG ATTTACCTGA TAAATTTAAT GCTTATTTAG CAAATAAATG GGTTTTTATA  
T I T N D R L S S A N L Y I N G V L M G  
541 ACTATTACTA ATGATAGATT ATCTTCTGCT AATTTGTATA TAAATGGAGT ACTTATGGGA  
S A E I T G L G A I R E D N N I T L K L  
601 AGTGCAGAAA TTAAGGTTT AGGAGCTATT AGAGAGGATA ATAATATAAC ATTAAGGATA  
D R C N N N N Q Y V S I D K F R I F C K  
661 GATAGATGTA ATAATAATAA TCAATACGTT TCTATTGATA AATTTAGGAT ATTTTGCAAA  
A L N P K E I E K L Y T S Y L S I T F L  
721 GCATTAAATC CAAAAGAGAT TGAAAAATTA TACACAAGTT ATTTATCTAT AACCTTTTTA  
R D F W G N P L R Y D T D R \*(SEQ ID NO:48)-  
781 AGAGACTTCT GGGGAAACCC TTTACGATAT GATACAGATA GGTAG (SEQ ID NO:47)-

FIG.29

55/62

```

1   ATGGAGTCTC CCTCGGCCCC TCCCCACAGA TGGTGCATCC CCTGGCAGAG GCTCCTGCTC
61  ACAGCCTCAC TTCTAACCTT CTGGAACCCG CCCACCACTG CCAAGCTCAC TATTGAATCC
121 ACGCCGTTCA ATGTCGCAGA GGGGAAGGAG GTGCTTCTAC TTGTCCACAA TCTGCCCCAG
181 CATCTTTTTG GCTACAGCTG GTACAAAGGT GAAAGAGTGG ATGGCAACCG TCAAATTATA
241 GGATATGTAA TAGGAACTCA ACAAGCTACC CCAGGGCCCCG CATAAGTGG TCGAGAGATA
301 ATATACCCCA ATGCATCCCT GCTGATCCAG AACATCATCC AGAATGACAC AGGATTCTAC
361 ACCCTACACG TCATAAAGTC AGATCTTGTG AATGAAGAAG CAACTGGCCA GTTCCGGGTA
421 TACCCGGAGC TGCCCAAGCC CTCCATCTCC AGCAACAACCT CCAAACCCGT GGAGGACAAG
481 GATGCTGTGG CCTTCACCTG TGAACCTGAG ACTCAGGACG CAACCTACCT GTGGTGGGTA
541 AACAATCAGA GCCTCCCGGT CAGTCCCAGG CTGCAGCTGT CCAATGGCAA CAGGACCCTC
601 ACTCTATTCA ATGTCACAAG AAATGACACA GCAAGCTACA AATGTGAAAC CCAGAACCCA
661 GTGAGTGCCA GCGCGAGTGA TTCAGTCATC CTGAATGTCC TCTATGGCCC GGATGCCCCC
721 ACCATTTCCC CTCTAAACAC ATCTTACAGA TCAGGGGAAA ATCTGAACCT CTCCTGCCAC
781 GCAGCCTCTA ACCCACCTGC ACAGTACTCT TGGTTTGTCA ATGGGACTTT CCAGCAATCC
841 ACCCAAGAGC TCTTTATCCC CAACATCACT GTGAATAATA GTGGATCCTA TACGTGCCAA
901 GCCCATAACT CAGACACTGG CCTCAATAGG ACCACAGTCA CGACGATCAC AGTCTATGCA
961 GAGCCACCCA AACCCTTCAT CACCAGCAAC AACTCCAACC CCGTGGAGGA TGAGGATGCT
1021 GTAGCCTTAA CCTGTGAACC TGAGATTGAG AACACAACCT ACCTGTGGTG GGTAAATAAT
1081 CAGAGCCTCC CGGTCAGTCC CAGGCTGCAG CTGTCCAATG ACAACAGGAC CCTCACTCTA
1141 CTCAGTGTCA CAAGGAATGA TGTAGGACCC TATGAGTGTG GAATCCAGAA CGAATTAAGT
1201 GTTGACCACA GCGACCCAGT CATCCTGAAT GTCCTCTATG GCCCAGACGA CCCCACCATT
1261 TCCCCCTCAT ACACCTATTA CCGTCCAGGG GTGAACCTCA GCCTCTCCTG CCATGCAGCC
1321 TCTAACCCAC CTGCACAGTA TTCTTGGCTG ATTGATGGGA ACATCCAGCA ACACACACAA
1381 GAGCTCTTTA TCTCCAACAT CACTGAGAAG AACAGCGGAC TCTATACCTG CCAGGCCAAT
1441 AACTCAGCCA GTGGCCACAG CAGGACTACA GTCAAGACAA TCACAGTCTC TGCAGGAGCTG
1501 CCCAAGCCCT CCATCTCCAG CAACAACCTC AAACCCGTGG AGGACAAGGA TGCTGTGGCC
1561 TTCACCTGTG AACCTGAGGC TCAGAACACA ACCTACCTGT GGTGGGTAAG TGCTCAGAGC
1621 CTCCCAGTCA GTCCCAGGCT GCAGCTGTCC AATGGCAACA GGACCCTCAC TCTATTCAAT
1681 GTCACAAGAA ATGACGCAAG AGCCTATGTA TGTGGAATCC AGAACTCAGT GAGTGCAAAC
1741 CGCAGTGACC CAGTCACCCT GGATGTCCTC TATGGGCCGG ACACCCCAT CATTTCCCCC
1801 CCAGACTCGT CTTACCTTTC GGGAGCGAAC CTCAACCTCT CCTGCCACTC GGCCTCTAAC
1861 CCATCCCCGC AGTATTGTTG GCGTATCAAT GGGATACCGC AGCAACACAC ACAAGTTCTC
1921 TTTATCGCCA AAATCACGCC AAATAATAAC GGGACCTATG CCTGTTTTGT CTCTAACTTG
1981 GCTACTGGCC GCAATAATTC CATAGTCAAG AGCATCACAG TCTCTGCATC TGGAACTCTA
2041 GATTCAACAC CAATTCCATT TTCTTATTCT AAAAATCTGG ATTGTTGGGT TGATAATGAA
2101 GAAGATATAG ATGTTATATT AAAAAAGAGT ACAATTTTAA ATTTAGATAT TAATAATGAT
2161 ATTATATCAG ATATATCTGG GTTTAATTCA TCTGTAATAA CATATCCAGA TGCTCAATTG
2221 GTGCCCGGAA TAAATGGCAA AGCAATACAT TTAGTAAACA ATGAATCTTC TGAAGTTATA
2281 GTGCATAAAG CTATGGATAT TGAATATAAT GATATGTTTA ATAATTTTAC CGTTAGCTTT
2341 TGGTTGAGGG TTCCTAAAGT ATCTGCTAGT CATTTAGAAC AATATGGCAC AAATGAGTAT
2401 TCAATAATTA GCTCTATGAA AAAACATAGT CTATCAATAG GATCTGGTTG GAGTGTATCA

```

FIG.30A-1

56/62

2461 CTTAAAGGTA ATAACTTAAT ATGGACTTTA AAAGATTCCG CGGGAGAAGT TAGACAAATA  
2521 ACTTTTAGGG ATTTACCTGA TAAATTTAAT GCTTATTTAG CAAATAAATG GGTTTTTATA  
2581 ACTATTACTA ATGATAGATT ATCTTCTGCT AATTTGTATA TAAATGGAGT ACTTATGGGA  
2641 AGTGCAGAAA TTA CTGGTTT AGGAGCTATT AGAGAGGATA ATAATATAAC ATTAAACTA  
2701 GATAGATGTA ATAATAATAA TCAATACGTT TCTATTGATA AATTTAGGAT ATTTTGCAAA  
2761 GCATTAAATC CAAAAGAGAT TGAAAAATTA TACACAAGTT ATTTATCTAT AACCTTTTTA  
2821 AGAGACTTCT GGGGAAACCC TTTACGATAT GATATAG (SEQ ID NO:49)

FIG.30A-2



57/62

```

1   ATGGGCAGCC CCAGCGCCCC CCTGCACCGC TGGTGCATCC CCTGGCAGAC
    CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG
101 CCCAGCTGAC CATCGAGAGC CGCCCCTTCA ACGTGGCCGA GGGCAAGGAG
    GTGCTGCTGC TGGCCCACAA CGTGAGCCAG AACCTGTTCTG GCTACATCTG
201 GTACAAGGGC GAGCGCGTGG ACGCCAGCCG CCGCATCGGC AGCTGCGTGA
    TCCGCACCCA GCAGATCACC CCCGGCCCCG CCCACAGCGG CCGCGAGACC
301 ATCGACTTCA ACGCCAGCCT GCTGATCCAC AACGTGACCC AGAGCGACAC
    CGGCAGCTAC ACCATCCAGG TGATCAAGGA GGACCTGGTG AACGAGGAGG
401 CCACCGGCCA GTTCCGCGTG TACCCGAGC TGCCCAAGCC CTACATCAGC
    AGCAACAACA GCAACCCCGT GGAGGACAAG GACGCCGTGG CCCTGACCTG
501 CGAGCCCGAG ACCCAGGACA CCACCTACCT GTGGTGGGTG AACAACCAGA
    GCCTGCCCCG GAGCCCCCGC CTGGAGCTGA GCAGCGACAA CCGCACCCCTG
601 ACCGTGTTCA ACATCCCCCG CAACGACACC ACCAGCTACA AGTGCGAGAC
    CCAGAACCCC GTGAGCGTGC GCCGCAGCGA CCCCCTGACC CTGAACGTGC
701 TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CCCCTACCGC
    GCCGGCGAGA ACCTGAACCT GACCTGCCAC GCCGCCAGCA ACCCCACCGC
801 CCAGTACTTC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC
    TGTTTCATCC CAACATCACC GTGAACAACA GCGGCAGCTA CATGTGCCAG
901 GCCCACAACA GCGCCACCGG CCTGAACCGC ACCACCGTGA CCGCCATCAC
    CGTGTACGCC GAGCTGCCCC AGCCCTACAT CACCAGCAAC AACAGCAACC
1001 CCATCGAGGA CAAGGACGCC GTGACCCTGA CCTGCGAGCC CGAGACCCAG
    GACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGA GCGTGAGCAG
1101 CCGCCTGGAG CTGAGCAACG ACAACCGCAC CCTGACCGTG TTCAACATCC
    CCCGCAACGA CACCACCTTC TACGAGTGCG AGACCCAGAA CCCCCTGAGC
1201 GTGCGCCGCA GCGACCCCGT GACCCTGAAC GTGCTGTACG GCCCCGACGC
    CCCCACCATC AGCCCCCTGA ACACCCCTA CCGCGCCGGC GAGAACCTGA
1301 ACCTGAGCTG CCACGCCGCC AGCAACCCCG CCGCCAGTA CAGCTGGTTC
    GTGAACGGCA CCTTCCAGCA GAGCACCCAG GAGCTGTTCA TCCCCAACAT
1401 CACCGTGAAC AACAGCGGCA GCTACATGTG CCAGGCCAC AACAGCGCCA
    CCGGCCTGAA CCGCACCACC GTGACCGCCA TCACCGTGTA CGTGGAGCTG
1501 CCCAAGCCCT ACATCAGCAG CAACAACAGC AACCCCATCG AGGACAAGGA
    CGCCGTGACC CTGACCTGCG AGCCCGTGGC CGAGAACACC ACCTACCTGT
1601 GGTGGGTGAA CAACCAGAGC CTGAGCGTGA GCCCCGCTT GCAGCTGAGC
    AACGGCAACC GCATCCTGAC CCTGCTGAGC GTGACCCGCA ACGACACCGG
1701 CCCCTACGAG TCGGGCATCC AGAACAGCGA GAGCGCCAAG CGCAGCGACC
    CCGTGACCCCT GAACGTGACC TACGGCCCCG ACACCCCAT CATCAGCCCC
1801 CCCGACCTGA GCTACCGCAG CGGCGCCAAC CTGAACCTGA GCTGCCACAG
    CGACAGCAAC CCCAGCCCCC AGTACAGCTG GCTGATCAAC GGCACCCTGC
1901 GCCAGCACAC CCAGGTGCTG TTCATCAGCA AGATCACCAG CAACAACAGC
    GGCGCCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
2001 CATCGTGAAG AACATCAGCG TGAGCAGCGG CGACAGCTCT AGAAGCACCC
    CCATCCCATT CAGCTACAGC AAGAACCTGG ACTGCTGGGT GGACAACGAG

```

FIG.31A-1

58/62

2101 GAGGACATCG ACGTGATCCT GAAGAAGAGC ACCATCCTGA ACCTGGACAT  
CAACAACGAC ATCATCAGCG ACATCAGCGG CTTCAACAGC AGCGTGATCA  
2201 CCTACCCCGA CGCCCAGCTG GTGCCC GGCA TCAACGGCAA GGCCATCCAC  
CTGGTGAACA ACGAGAGCAG CGAGGTGATC GTGCACAAGG CCATGGACAT  
2301 CGAGTACAAC GACATGTTCA ACAACTTCAC CGTGAGCTTC TGGCTGAGAG  
TGCCTAAGGT GAGCGCCAGC CACCTGGAGC AGTACGGCAC CAACGAGTAC  
2401 AGCATCATCA GCAGCATGAA GAAGCACAGC CTGAGCATCG GCAGCGGCTG  
GAGCGTGAGC CTGAAGGGCA ACAACCTCAT CTGGACCCTG AAGGATAGCG  
2501 CCGGAGAGGT GAGACAGATC ACCTTCAGAG ACCTGCCCCA CAAGTTCAAT  
GCCTACCTGG CCAACAAGTG GGTGTTTCATC ACCATCACCA ACGACAGACT  
2601 GAGCAGCGCC AACCTGTACA TCAACGGCGT GCTCATGGGC AGCGCCGAGA  
TCACCGGCCT GGGCGCCATC AGAGAGGACA ACAACATCAC CCTGAAGCTG  
2701 GACAGATGCA ACAACAACAA CCAGTACGTG AGCATCGACA AGTTCCGGAT  
CTTCTGCAAG GCCCTGAACC CCAAGGAGAT CGAGAAGCTG TACACCAGCT  
2801 ACCTGAGCAT CACCTTCCTG AGAGACTTCT GGGGCAACCC CCTGAGATAC  
GACACCTAG (SEQ ID NO:50)

FIG.31A-2

59/62

```
1  MGSPSAPLHR WCIPWQTLLE TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWKYK ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELPKPYIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF WFNVTGTFQQS TQELFIPNIT VNNSGSYMCQ
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETQ
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRNDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQGSTQ ELFIPNITVN NSGSYMCQAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPEY CGIQNSSESAK RSDPVTLNVT YGPDTPIIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSS RSTPIPFYSY KNLDCWVDNE
701 EDIDVILKKS TILNLDINND IISDISGFNS SVITYPDAQL VPGINGKAIH
751 LVNNESESEVI VHKAMDIEYN DMFNFTVVSF WLRVPKVSAS HLEQYGTNEY
801 SIISSMKKHS LSIGSGWSVS LKGNNLIWTL KDSAGEVRQI TFRDLPDKFN
851 AYLANKWVFI TITNDRLLSA NLYINGVLMG SAEITGLGAI REDNNITLKL
901 DRCNNNNQYV SIDKFRIFCK ALNPKEIEKL YTSYLSITFL RDFWGNPLRY
951 DT* (SEQ ID NO:51)
```

FIG.31B

60/62

```

1   ATGGGCAGCC CCAGCGCCCC CCTGCACCGC TGGTGCATCC CCTGGCAGAC
    CCTGCTGCTG ACCGCCAGCC TGCTGACCTT CTGGAACCCC CCCACCACCG
101 CCCAGCTGAC CATCGAGAGC CGCCCCTTCA ACGTGGCCGA GGGCAAGGAG
    GTGCTGCTGC TGGCCCACAA CGTGAGCCAG AACCTGTTCG GCTACATCTG
201 GTACAAGGGC GAGCGCGTGG ACGCCAGCCG CCGCATCGGC AGCTGCGTGA
    TCCGCACCCA GCAGATCACC CCCGGCCCCG CCCACAGCGG CCGCGAGACC
301 ATCGACTTCA ACGCCAGCCT GCTGATCCAC AACGTGACCC AGAGCGACAC
    CGGCAGCTAC ACCATCCAGG TGATCAAGGA GGACCTGGTG AACGAGGAGG
401 CCACCGGCCA GTTCCGCGTG TACCCCGAGC TGCCCAAGCC CTACATCAGC
    AGCAACAACA GCAACCCCGT GGAGGACAAG GACGCCGTGG CCCTGACCTG
501 CGAGCCCGAG ACCCAGGACA CCACCTACCT GTGGTGGGTG AACAAACCAGA
    GCCTGCCCCG GAGCCCCCGC CTGGAGCTGA GCAGCGACAA CCGCACCTTG
601 ACCGTGTTCA ACATCCCCCG CAACGACACC ACCAGCTACA AGTGCGAGAC
    CCAGAACCCC GTGAGCGTGC GCCGCAGCGA CCCCCTGACC CTGAACGTGC
701 TGTACGGCCC CGACGCCCCC ACCATCAGCC CCCTGAACAC CCCCTACCGC
    GCCGGCGAGA ACCTGAACCT GACCTGCCAC GCCGCCAGCA ACCCCACCGC
801 CCAGTACTTC TGGTTCGTGA ACGGCACCTT CCAGCAGAGC ACCCAGGAGC
    TGTTTCATCC CAACATCACC GTGAACAACA GCGGCAGCTA CATGTGCCAG
901 GCCCACAACA GCGCCACCGG CCTGAACCGC ACCACCGTGA CCGCCATCAC
    CGTGTACGCC GAGCTGCCCA AGCCCTACAT CACCAGCAAC AACAGCAACC
1001 CCATCGAGGA CAAGGACGCC GTGACCCTGA CCTGCGAGCC CGAGACCCAG
    GACACCACCT ACCTGTGGTG GGTGAACAAC CAGAGCCTGA GCGTGAGCAG
1101 CCGCCTGGAG CTGAGCAACG ACAACCGCAC CCTGACCGTG TTCAACATCC
    CCCGCAACGA CACCACCTTC TACGAGTGCG AGACCCAGAA CCCCCTGAGC
1201 GTGCGCCGCA GCGACCCCGT GACCCTGAAC GTGCTGTACG GCCCCGACGC
    CCCCACCATC AGCCCCCTGA ACACCCCTA CCGCGCCGGC GAGAACCTGA
1301 ACCTGAGCTG CCACGCCGCC AGCAACCCCG CCGCCCAGTA CAGCTGGTTC
    GTGAACGGCA CCTTCCAGCA GAGCACCCAG GAGCTGTTCA TCCCCAACAT
1401 CACCGTGAAC AACAGCGGCA GCTACATGTG CCAGGCCAC AACAGCGCCA
    CCGGCCTGAA CCGCACCACC GTGACCGCCA TCACCGTGTA GGTGGAGCTG
1501 CCCAAGCCCT ACATCAGCAG CAACAACAGC AACCCCATCG AGGAAEAGGA
    CGCCGTGACC CTGACCTGCG AGCCCGTGGC CGAGAACACC ACCTACCTGT
1601 GGTGGGTGAA CAACCAGAGC CTGAGCGTGA GCCCCCGCCT GCAGCTGAGC
    AACGGCAACC GCATCCTGAC CCTGCTGAGC GTGACCCGCA ACGACACCGG
1701 CCCCTACGAG TCGGGCATCC AGAACAGCGA GAGCGCCAAG CGCAGCGACC
    CCGTGACCCT GAACGTGACC TACGGCCCCG ACACCCCAT CATCAGCCCC
1801 CCGGACCTGA GCTACCGCAG CGGCGCCAAC CTGAACCTGA GCTGECACAG
    CGACAGCAAC CCCAGCCCCC AGTACAGCTG GCTGATCAAC GGCACCCTGC
1901 GCCAGCACAC CCAGGTGCTG TTCATCAGCA AGATCACCAG CAACAACAGC
    GCGGCCTACG CCTGCTTCGT GAGCAACCTG GCCACCGGCC GCAACAACAG
2001 CATCGTGAAG AACATCAGCG TGAGCAGCGG CGACAGCTCT AGAACCCCTC
    AGAACATCAC CGATCTGTGC GCCGAGTACC ACAACACCCA GATCTACACC

```

FIG.32A-1

61/62

2101 CTGAACGACA AGATCTTCAG CTACACCGAG AGCCTGGCCG GCAAGAGAGA  
GATGGCCATC ATCACCTTCA AGAACGGCGC CATCTTCCAG GTGGAGGTGC  
2201 CCGGCAGCCA GCACATCGAC AGCCAGAAGA AGGCCATCGA GCGGATGAAG  
GACACCCTGC GGATCGCCTA CCTCACCGAG GCCAAGGTGG AGAAGCTGTG  
2301 CGTGTGGAAC AACAAGACCC CTCACGCCAT CGCCGCCATC AGCATGGCCA  
ATTGATAAG (SEQ ID NO:52)

FIG.32A-2

62/62

```
1   MGSPSAPLHR WCIPWQTL L TASLLTFWNP PTTAQLTIES RPFNVAEGKE
51  VLLLAHNVSQ NLFGYIWYKG ERVDASRRIG SCVIRTQQIT PGPAHSGRET
101 IDFNASLLIH NVTQSDTGSY TIQVIKEDLV NEEATGQFRV YPELKPYPIS
151 SNNSNPVEDK DAVALTCEPE TQDTTYLWWV NNQSLPVSPR LELSSDNRTL
201 TVFNIPRNDT TSYKCETQNP VSVRRSDPVT LNVLYGPDAP TISPLNTPYR
251 AGENLNLTCH AASNPTAQYF W FVNGTFQQS TQELFIPNIT VNNSGSYMCQ
301 AHNSATGLNR TTVTAITVYA ELPKPYITSN NSNPIEDKDA VTLTCEPETO
351 DTTYLWWVNN QSLSVSSRLE LSNDNRTLTV FNIPRIKDTTF YECETQNPVS
401 VRRSDPVTLN VLYGPDAPTI SPLNTPYRAG ENLNLSCHAA SNPAAQYSWF
451 VNGTFQQSTQ ELFIPNITVN NSGSYMCQAH NSATGLNRTT VTAITVYVEL
501 PKPYISSNNS NPIEDKDAVT LTCEPVAENT TYLWWVNNQS LSVSPRLQLS
551 NGNRILTLLS VTRNDTGPEY CGIQNSESAK RSDPVTLNVT YGPDTPIIISP
601 PDLSYRSGAN LNLSCHSDSN PSPQYSWLIN GTLRQHTQVL FISKITSNNS
651 GAYACFVSNL ATGRNNSIVK NISVSSGDSS RTPQNITDLC AEYHNTQIYT
701 LNDKIFSYTE SLAGKREMAI ITFKNGAIFQ VEVPGSQHID SQKKAIERMK
751 DTLRIAYLTE AKVEKLCVWN NKTPHAIAAI SMAN** (SEQ ID NO:53)
```

FIG.32B